

UNIVERSITE DE NICE-SOPHIA ANTIPOlis - UFR Sciences

Ecole Doctorale des Sciences Fondamentales et Appliquées

T H E S E

pour obtenir le titre de

Docteur en Sciences

de l'UNIVERSITE de Nice-Sophia Antipolis

Discipline : Chimie

présentée et soutenue par

Claire de MARCH

**MODELISATION DES MECANISMES
MOLECULAIRES
DE LA PERCEPTION DES ODEURS**

Thèse dirigée par **Jérôme GOLEBIOWSKI**

soutenue le 23 octobre 2015

Jury :

Dr. Loïc BRIAND	CSGA, Dijon	Rapporteur
Pr. Bernard OFFMANN	Université de Nantes	Rapporteur
Dr. Moustafa BENSAFI	Université Claude Bernard, Lyon	Examinateur
Pr. Xavier FERNANDEZ	Université de Nice-Sophia Antipolis	Examinateur
Dr. Gilles SICARD	Université Aix-Marseille	Examinateur
Pr. Jérôme GOLEBIOWSKI	Université de Nice-Sophia Antipolis	Directeur de thèse

Remerciements

Mes travaux de thèses ont été réalisés dans l'équipe Arôme, Parfum, Synthèse et Modélisation de l'Institut de Chimie de Nice sous la responsabilité du Professeur Jérôme Golebiowski. Je tiens à le remercier chaleureusement pour sa disponibilité et son implication durant ces trois années. M'autorisant à dépasser mon rôle d'étudiante et étant toujours d'un grand soutien, il a su me faire évoluer vers la chercheuse que je désirerais être et a largement contribué au succès de cette thèse.

Je remercie également la fondation Edmond Roudnitska et son conseil d'administration, Michel Roudnitska, Catherine Rouby, Maurice Chastrette, Anne-Marie Mouly et André Holley pour m'avoir fait confiance il y a trois ans en finançant ce doctorat. Plus particulièrement, je tiens à remercier vivement le docteur Gilles Sicard pour avoir soutenu ce projet. Il a veillé au bon déroulement de mes travaux de recherche tout au long de cette thèse en m'apportant une partie de la richesse de ses connaissances en perception des odeurs. Je lui en suis très reconnaissante.

Je tiens à remercier les rapporteurs de cette thèse, Dr. Loïc Briand et Pr. Bernard Offmann pour avoir accepté de l'évaluer. Ils sont de grands experts dans leur domaine respectif et me font l'honneur d'être associés à ce travail. Merci au Pr. Xavier Fernandez, avec qui j'ai débuté mon expérience à l'ICN, de finir cette aventure avec moi en acceptant d'examiner ces travaux. Je remercie Dr. Moustafa Bensafi d'avoir accepté mon invitation à être membre du jury. Je tenais à sa présence et je suis heureuse qu'il participe au jury de cette thèse.

Un grand merci à tous nos collaborateurs, Dr. Nicolas Baldovini, Dr. Anne-Marie Le Bon, Dr. Minghong Ma de l'université de Pennsylvanie, Dr. Hiroaki Matsunami de l'université Duke, Pr. Cheil Moon et SangEun Ryu du DGIST pour leurs données expérimentales et l'initiation de nombreux projets.

Je n'oublie pas l'équipe de modélisation avec qui j'ai passé la plupart de mon temps et qui aura été une sorte de famille adoptive à Nice. Au patriarche, Serge, merci d'avoir tenté à tes dépens de faire de moi une cuisinière alsacienne. Je repars avec la règle des trois : beurre/crème fraîche/lardon.

Merci à Tonton Fabien, toujours présent pour discuter claquettes et politique. Seb, merci pour ta bonne humeur, ta convivialité et tes conseils sur le docking. Merci à Martine, mon alliée féminine. Un remerciement particulier à l'un de ses membres, Julien Diharce, mon grand frère, qui ne m'a jamais laissée sombrer devant certaines de mes lacunes et a toujours su me proposer une bière au bon moment. Et Jean-Baptiste, le 'ptit' dernier, merci d'avoir supporté ta peste de grande sœur. Ton arrivée a été une bouffée d'air frais.

Et bien entendu, un grand merci aux lards et aux poneys.

Enfin, j'ai une pensée particulière pour ma famille qui était si loin pendant cette thèse et pourtant si présente. Merci pour leur soutien et leur enthousiasme qui m'ont permis de souffler et de tenir le coup durant ces trois années.

Table des matières

INTRODUCTION.....	1
PARTIE 1 : LES MECANISMES DE LA PERCEPTION DES ODEURS.....	7
Article 1 - Vers l'étude des Relations Structure-Odeur à l'ère post génomique.....	9
Article 2 - Les étapes moléculaires de la perception des odeurs décrites par les approches de modélisation moléculaire.....	51
PARTIE 2 : MODELISATION MOLECULAIRE DES RECEPTEURS OLFACTIFS.....	77
Articles 3 et 4 - La modélisation moléculaire des ROs : de la séquence à la structure.....	79
PARTIE 3 : RELATIONS STRUCTURE-FONCTION DES RECEPTEURS OLFACTIFS.....	107
Articles 5 et 6 - Lien olfactophore-code combinatoire de ROs.....	109
Article 7 - Le calcul de l'affinité odorant/récepteur discrimine les agonistes des non-agonistes de hOR1G1.....	145
Article 8 - Le spectre de reconnaissance d'un RO est modulé par son affinité avec les odorants mais aussi par sa barrière d'activation.....	157
Articles 9 et 10 - Activation des récepteurs - vers la déorphanisation computationnelle.....	179
CONCLUSIONS ET PERSPECTIVES.....	207
ANNEXE.....	213

Introduction

Introduction

Pourquoi la molécule de (+)-carvone déclenche-t-elle une odeur de carvi alors que son énantiomère, la (-)-carvone, évoque la menthe verte ? Pourquoi le musc cétoné et l'androsténol, qui présentent deux structures chimiques radicalement dissemblables, provoquent tous deux une odeur musquée ? La réponse à ces questions de chimiste nécessite d'aller au-delà des frontières des disciplines. Elle se cache dans la façon dont notre cerveau, dans sa fascinante complexité, traite l'information moléculaire portée par ces odorants. Notre perception des odeurs résulte de l'interaction entre une molécule et les neurones olfactifs situés à l'extrémité supérieure de notre cavité nasale. Afin d'être en mesure de discriminer un nombre spectaculaire de molécules odorantes, notre cerveau a établi une stratégie qui repose sur un code dit « combinatoire » d'activation neuronale. Au niveau cellulaire, près de 400 sous-types de neurones olfactifs expriment chacun un seul type de protéine réceptrice : le récepteur olfactif.

Un modèle universel qui reliera la structure des molécules odorantes à une odeur reste - bien que possible en principe – à établir. A l'heure actuelle, la recherche de nouveaux composés odorants repose principalement sur la synthèse de dérivés de molécules odorantes connues et est très sujette à la sérendipité. Contrairement à la vision où le stimulus est bien caractérisé d'un point de vue physique et dont la perception repose sur trois types de capteurs, la complexité de l'espace odorant et la taille de notre répertoire de récepteurs olfactifs ont freiné notre compréhension du codage et du traitement des odeurs. Par exemple, il n'y a, à ce jour, aucun consensus sur les caractéristiques clés de la perception des odeurs comme le nombre d'odeurs que l'humain est capable de discriminer ou comment des fonctionnalités telles que la valence hédonique ou l'identification d'une note olfactive sont codées.

L'odorat ainsi que le goût sont les deux sens dits « chimiques » qui permettent à l'homme de détecter les molécules présentes dans son environnement. L'étape clé des mécanismes de la perception des odeurs est l'interaction de ces molécules avec nos neurones olfactifs. Ceux-ci expriment des protéines permettant de transformer les signaux chimiques portés par les molécules odorantes en influx neuronaux activant différentes zones de notre cerveau, qu'il interprétera comme une odeur. Linda Buck et Richard Axel qui ont été récompensés par le prix Nobel de médecine en 2004 pour leur découverte^[1] des Récepteurs Olfactifs (ROs), qui sont, à ce jour, décomptés au nombre de 396 chez l'homme.^[2] Lors de la perception d'une odeur, chacune de ces protéines interagit avec un odorant et contribue ainsi à un code combinatoire d'activation de récepteurs. (Figure 1) Cette « carte d'identité » est supposée être liée à l'odeur de cette molécule.

De plus, des phénomènes dits péri-récepteurs interviennent lors de l'inhalation de composés odorants. Ces phénomènes impliquent d'autres types de protéines ou d'enzymes, dont le rôle, le nombre ou les caractéristiques biologiques sont encore mal connus.^[3] Ces inconnues contribuent à

complexifier les premières études de relation entre la structure d'une molécule et son odeur par la difficulté de prendre en compte les protagonistes biologiques impliqués.

Malgré tout, sur le principe, on associe à chaque odeur son propre code combinatoire de récepteurs. Cette stratégie à 396 variables utilisée par notre cerveau est en accord avec l'extraordinaire nombre d'odeurs discriminable par l'homme.[4-6]

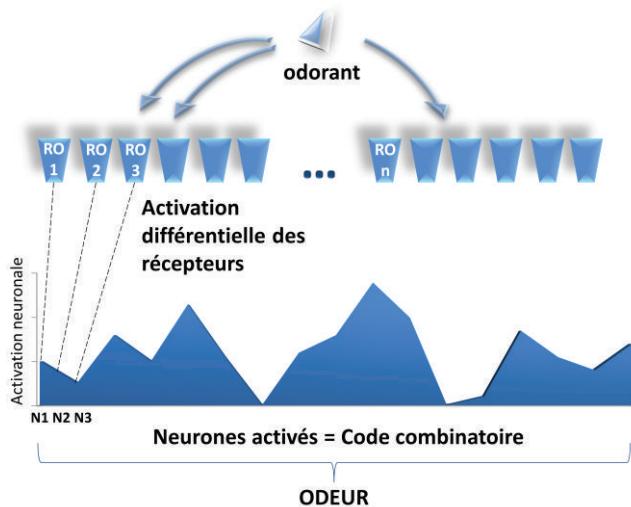


Figure 1. Principe de l'encodage des odeurs. Une molécule odorante interagit avec la totalité du répertoire de récepteurs olfactifs de façon différentielle. L'activation de ces récepteurs peut être reliée à l'activation du neurone porteur du RO. L'ensemble des différentes activations forme un 'code combinatoire' et constitue une carte d'identité de l'odorant, en principe associée à son odeur.

Les récepteurs olfactifs sont des protéines transmembranaires appartenant à la famille des Récepteurs Couplés aux Protéines G (RCPGs). Leur mécanisme d'activation s'appuie sur la favorisation d'un état actif par rapport à un état inactif lorsque qu'une molécule agoniste est détectée. Si quelques RCPGs ont été cristallisés, aucune structure de récepteur olfactif n'est à ce jour disponible. Cependant, depuis quelques années, les progrès en matière d'outils informatiques et de connaissances de la physique des atomes permettent l'utilisation de méthodes de modélisation moléculaire efficaces dans le cadre de l'étude théorique de systèmes biologiques. L'application de ces méthodes aux récepteurs olfactifs, guidée par des données expérimentales, en font un outil prédictif performant pour l'étude de leurs interactions avec les odorants. C'est majoritairement grâce à cette approche que les récepteurs olfactifs sont décrits dans ce document.

Ma démarche scientifique s'est basée sur une série de questions qu'il me paraissait crucial d'aborder afin de permettre dans un futur proche de décrire la relation entre une structure chimique et une odeur sur une base rationnelle et physiologiquement inspirée. Je me suis focalisée sur l'étude des récepteurs olfactifs et la compréhension de leur mécanisme de reconnaissance. Mais comment

obtenir un modèle prédictif de ces protéines, tant d'un point de vue structural que vis-à-vis de son interaction avec un odorant ? Quels paramètres régulent leurs interactions avec des composés odorants ? Comment prédire leur comportement dynamique ? Les travaux originaux réalisés lors de cette thèse ont pour objectif d'apporter des éléments de réponse à ces questions.

Dans un premier temps, l'état de l'art en matière d'étude des relations structure-odeur est présenté sous forme d'un article de revue qui souligne l'importance de la prise en compte des protagonistes biologiques de l'olfaction. Une revue plus technique, focalisée sur les différentes contributions de la modélisation moléculaire à l'étude de l'olfaction, est ensuite proposée. Ensuite, une série d'articles de recherche, pour la plupart multidisciplinaires, tente d'apporter des réponses fondamentales sur les mécanismes régissant le fonctionnement de nos récepteurs olfactifs et leurs interactions avec les odorants. Dans ce but, la modélisation moléculaire est associée à des techniques allant de la synthèse organique au génie génétique à travers des collaborations nationales et internationales.

Ma contribution peut se résumer de la manière suivante. Dans un premier temps j'ai établi un protocole de reconstruction de la structure tridimensionnelle des récepteurs olfactifs de mammifère. Ensuite j'ai identifié les bases de la relation entre la séquence d'un récepteur et son mécanisme d'activation en fonction de la structure d'une molécule odorante liée à sa cavité. J'ai pu montrer que l'analyse des structures de molécules d'une même famille olfactive pouvait conduire à l'identification des récepteurs impliqués dans leur perception. L'ensemble de ces résultats constitue les bases pour l'étude des relations structure-odeur à l'ère post génomique.

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] A. Matsui, Y. Go, Y. Niimura, Degeneration of Olfactory Receptor Gene Repertoires in Primates: No Direct Link to Full Trichromatic Vision, *Molecular Biology and Evolution* 27 (2010) 1192-1200.
- [3] J.-M. Heydel, A. Coelho, N. Thiebaud, A. Legendre, A.-M.L. Bon, P. Faure, F. Neiers, Y. Artur, J. Golebiowski, L. Briand, Odorant-Binding Proteins and Xenobiotic Metabolizing Enzymes: Implications in Olfactory Perireceptor Events, *The Anatomical Record* 296 (2013) 1333-1345.
- [4] C. Bushdid, M.O. Magnasco, L.B. Vosshall, A. Keller, Humans Can Discriminate More than 1 Trillion Olfactory Stimuli, *Science* 343 (2014) 1370-1372.
- [5] M. Meister, On the dimensionality of odor space, *eLife* 4 (2015) e07865.
- [6] R.C. Gerkin, J.B. Castro, The number of olfactory stimuli that humans can discriminate is still unknown, *eLife* 4 (2015) e08127.

Introduction

Les mécanismes de la perception des odeurs

Article 1 – Vers l'étude des Relations Structure-Odeur à l'ère post génomique

L'étude des relations structure-odeur a débuté dès la découverte des structures chimiques des molécules odorantes. Elle bénéficie depuis une vingtaine d'années de l'entrée des sciences du vivant dans l'ère post génomique. Sarah Richardson définit cette période en 2014 comme étant la « période suivant le séquençage complet du génome humain, qui a été dominée par la transdisciplinarité, la rapidité et la centralité des technologies informatiques qui marquent les sciences du vivant contemporaines. »

Dans le domaine de l'olfaction, cette ère commence en 1991. L. Buck et R. Axel découvrent dans le génome des mammifères les gènes codant pour les protéines impliquées dans notre perception des odeurs : les récepteurs olfactifs (ROs). Les communautés scientifiques engagées dans la recherche des relations structure-odeur englobent donc des domaines aussi variés que les neurosciences, la biochimie, la chimie organique, analytique, la bio-informatique, l'analyse sensorielle ou la linguistique. La communication entre ces domaines d'expertise est un réel défi mais est nécessaire pour regrouper toute la transdisciplinarité requise par les sciences de l'olfaction. Ici nous présentons les différents domaines des sciences dites « dures » et « humaines » qui paraissent incontournables dans l'étude des relations structure-odeur projetée à l'ère post génomique.

Qu'est-ce qu'une structure? Qu'est-ce qu'une odeur?

L'élucidation des relations existantes entre la structure d'une molécule et son odeur commence par la définition claire de ces deux termes. La structure est constituée de différents atomes liés entre eux par des liaisons chimiques formant ainsi une structure tridimensionnelle. La rotation autour de ces liaisons permet à la molécule de prendre différentes formes dans l'espace appelées conformations. L'odeur associée à une molécule reste, quant à elle, difficile à caractériser de façon qualitative et quantitative. Les fortes différences inter et intra culturelles et leur impact sur les termes employés rendent la description olfactive extrêmement variable d'un individu à l'autre. Poser des descripteurs olfactifs sur une molécule permet malgré tout de les catégoriser. A titre d'exemple, un composé peut posséder une odeur de rose et de miel, un autre de rose et d'herbe ; ces deux odorants appartiennent donc à la sous famille des notes dites « rosées ». On peut aussi les regrouper de façon plus large avec des molécules possédant des odeurs de lilas, jasmin ou lavande sous la large famille des notes florales. La recherche la plus directe de la relation entre une structure et une odeur consiste à trouver les similarités structurales entre les composés appartenant à la même famille olfactive. Cette approche permet de déterminer les propriétés moléculaires associées à une odeur

ciblée. Ces propriétés peuvent être, par exemple, de nature physico-chimique et être représentées sous la forme d'un modèle tridimensionnel. Ils sont appelés olfactophores (Figure1-a). Bien qu'efficaces dans certain cas, ces méthodes atteignent leurs limites quand les structures chimiques des molécules au sein d'une même famille olfactive sont très différentes (Figure1-b). On notera que ces approches sont uniquement basées sur la structure chimique des molécules et qu'à aucun moment le récepteur biologique n'est évoqué.

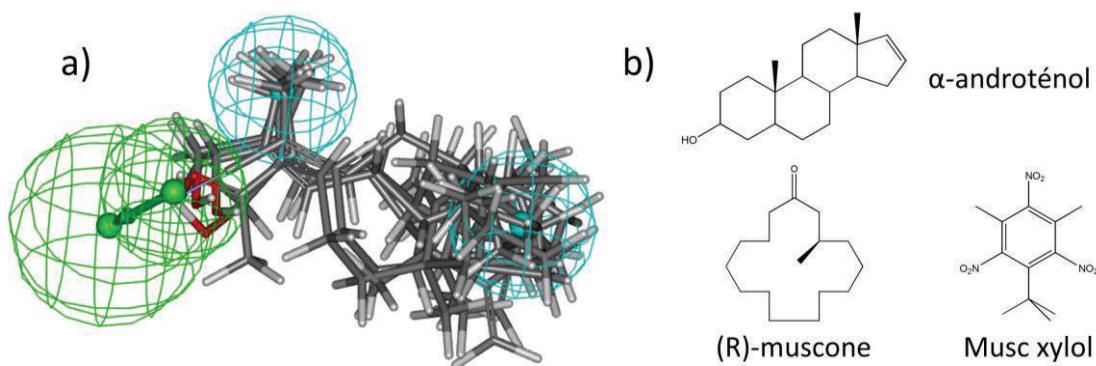


Figure 1. a) modèle tridimensionnel de l'olfactophore du bois de santal. Les sphères représentent les propriétés physico-chimiques nécessaires aux molécules à odeur santalée (accepteur de liaisons hydrogène en vert et parties hydrophobes en bleu). Les composés santalés (en représentation traits) sont superposés sur le modèle et correspondent aux trois critères. b) Exemples représentant les structures typiques des molécules à odeur musquée. Dans ce cas, l'approche olfactophore est très délicate.

Comment prendre en compte les ROs dans les relations structure-odeur ?

La complexité des mécanismes de l'olfaction est notamment liée à la variabilité des protagonistes biologiques. La subtilité de leur reconnaissance moléculaire doit être prise en compte. Il apparaît nécessaire d'étudier les interactions entre les molécules odorantes et nos ROs. Des expériences de biologie cellulaire permettent de mesurer l'activation d'un récepteur lors des stimulations par une molécule odorante *in vitro*. Ces expériences menées sur des cultures de cellules sont éloignées des conditions physiologiques de l'olfaction et la question peut se poser quant à la pertinence des résultats obtenus. Dans ce type de méthode, l'observation des récepteurs olfactifs se fait indirectement à travers la réponse de cellules qui sont déjà des systèmes complexes. Mais comment fonctionnent exactement ces machines moléculaires au niveau atomique ?

Les ROs, qui sont estimés au nombre de 396 chez l'humain, ont la propriété de s'activer pour un nombre important d'odorants. De plus, un même composé peut engendrer la réponse de plusieurs ROs. Le comportement d'un odorant face à la totalité des 396 ROs lui est propre et est appelé code combinatoire. Ce code de ROs activés semblerait pouvoir être relié à l'odeur du composé, supposant qu'un code combinatoire corresponde à une odeur. L'élucidation de ce code serait donc la clé de

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

l'étude des relations structure-odeur à l'ère post génomique. De nombreuses questions restent cependant en suspens. Au sein de ce code, certains ROs ont-ils plus de poids ? La réponse d'un RO face à un odorant se limite-t-elle à activation ou l'absence d'activation ou peut-elle être plus subtile ? Un RO peut-il être spécifique à une famille olfactive ? Le code combinatoire de la perception des odeurs est-il déchiffré ? Existe-t-il des méthodes alternatives à l'expérience *in vitro* pour obtenir ce code ?

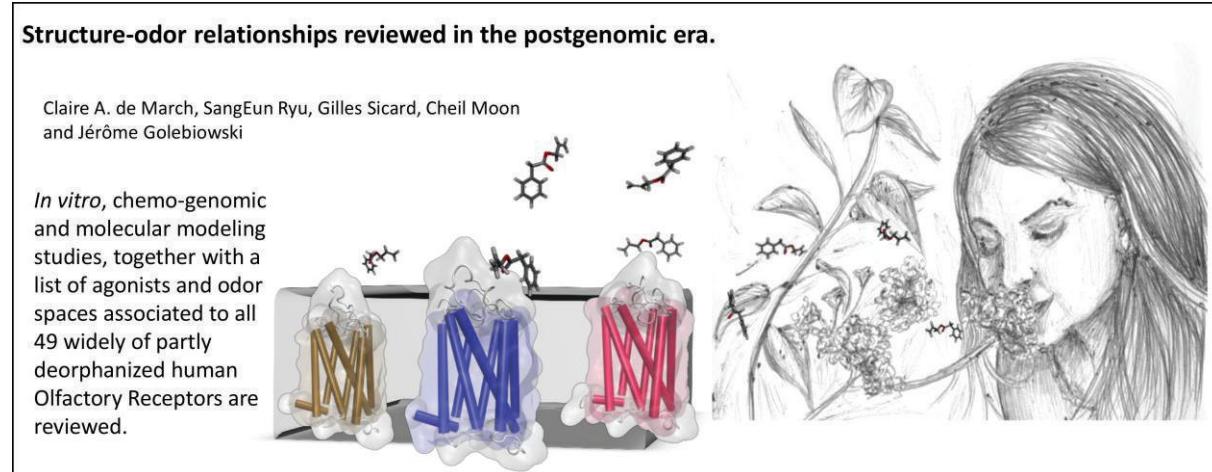
Dans cette revue, nous apportons des éléments de réponse à ces questions. Nous établissons également un début de lien entre une famille olfactive ou une molécule odorante et un code combinatoire de ROs grâce aux données expérimentales disponibles dans la littérature. Nous réalisons ainsi un premier pas vers l'établissement des relations structure-odeur projeté à l'ère post génomique.

Cet article a été écrit avec l'aide précieuse de nos collaborateurs Coréen du DGIST, SangEun Ryu et le Pr. Cheil Moon, concernant la partie « Cellular expression system for the study of ORs ». Le Dr. Gilles Sicard nous a également apporté son expertise en identifiant notamment les points délicats à surmonter dans l'établissement des relations structure-odeur.

Article 1:

Structure-odor relationships reviewed in the postgenomic era.

Claire A. de March, SangEun Ryu, Gilles Sicard, Cheil Moon, Jérôme Golebiowski, FFJ 30 (2015) 342-361



"But, side by side and inside this spiritual love I have for you there is also a wild beast-like craving for every inch of your body, for every secret and shameful part of it, for every odour and act of it."

James Joyce, in a letter to Nora Barnacle. December 2, 1909

Keywords: Olfactory receptor, odorant, deorphanization, smell, interaction

Abstract

This review reports knowledge about odorant-olfactory receptor interactions and their projection within the field of structure-odor relationships. We provide a list of agonists and odor spaces associated with deorphanized human olfactory receptors. The link between olfactory receptor responses, differential perception and subtleties within odor families is discussed.

Introduction

The perception of an odor is the result of an extraordinary complex cascade of events. The very first steps of this perception involve the interaction of chemicals with our olfactory neurons. At the molecular level, our neurons express proteins that play a role in transforming this chemical signal into electrophysiological messages that are processed in the brain as an odor.[1] However, in parallel with physiological and genetic studies, chemists have tried early on to link chemical structures with odors.[2]

Based on the idea that the odor of a chemical is fully encoded within its structure, various structure-odor relationships (SOR) have been tentatively established. Some have shown limited success, whereas others, when focused on the chemical structures within a well-defined family, have offered some predictive models.[3]

Odor is a property that depends on several highly variable factors. The first difficulty lies in the definition of the class (or category) of smell, which is a prerequisite to establish SOR. Categories are named by general descriptors because they help individuals create a consensus around the description of odors. These descriptors are mostly subjective and therefore are not universal. Dravnieks (1985) proposed a list of more than one hundred descriptors for 144 odorants.[4] By analyzing a collection of ~2500 odor descriptions[5], Chastrette et al. (1988)[6] showed that only 3% of the used descriptors could lead to a fruitful structure-odor analysis. In fact, not all descriptors are associated to the same semantic level, and they do not describe the same properties of olfactory perception.[7] Some refer to classes of objects with associated smells ("floral", "fruity"), whereas others are associated at a chemical level ('camphoraceous', 'aldehydic', 'acidic'). In those latter sets, the molecules that generally share sensory descriptors also present related chemical features. For example, many linear aldehydes share a so-called 'aldehydic' olfactory note. In other categories that share the same olfactory note, the associated chemical structures can be so different that it is difficult to believe that they involve the same mechanism. Prototypical of this is the intriguing case of musky odors because these molecules encompass four different types of chemical structures: a macrocyclic structure as it is the case for muscone (a 15-membered ring); a polycyclic structure as galaxolide; a derivative of trinitrotoluene (*e.g.*, the musk ketone)[8]; or an aliphatic structure.[9] The

receptor that is involved in the recognition of muscone was recently shown to be rather specific to this musky series.[10, 11] Conversely, similar structures can elicit different smells, as in the case of β -santalol, which has a sandalwood odor that can be abolished in derivatives with extremely small modifications.[12] Other examples are gathered in table 1. Even more counterintuitively, enantiomers of carvone can elicit either ‘mint’ or ‘carvi’ smells depending on the configuration of the asymmetric carbon atom, as illustrated in a series of enantiomers shown in table 2.[13]

Table 1. Compounds with similar structures eliciting different smells.[8]		
coconut	mint	herbaceous
soup	caramel	floral
spicy		

Table 2. Enantiomers associated to different smells.[13]

(S)-2-methylbutan-1-ol ethereal	(R)-2-methylbutan-1-ol waxy, fermented
(S)-limonene turpentine	(R)-limonene citrus
(S)-3-methylthio-hexanol spicy	(R)-3-methylthio-hexanol fruity, exotic
(S)-1-octen-3-ol moldy	(R)-1-octen-3-ol mushroom

These subtleties are rooted in the combinatorial code of olfactory receptor (OR) activation. After entering the nasal cavity, an odorant molecule can interact with various biological agents, each with a specific role involved in the detection of the odor signal, which is further transmitted to integrative neuronal networks. Each olfactory neuron possesses a single type of olfactory receptor allele that is embedded within its membrane [14] and interacts with odorant molecules (Figure 1). As each neuron only expresses one OR allele,[15] OR activation is equivalent, in principle, to the behavior of a stimulated neuron. Following the discovery of olfactory receptors in rats in 1991[1] and more importantly given the number of olfactory genes within the human genome, it becomes obvious that the strategy used by our brain to represent odorous chemical signals, *i.e.*, a combinatorial multineuronal code, cannot be dismissed.[1, 16] This confirms the seminal work of Polak published two decades before.[17] Over 900 ORs genes and pseudogenes were identified in 2001[18], but this number was revised to ~400 potentially functional ORs in 2010.[19] The primary rules governing an olfactory message are as follows: a receptor can interact with several molecules with potentially different structures, whereas a single odorant can interact with various receptors.[20] In this post-genomic era, the amount of information and knowledge about inter-individual variability makes the deciphering of the sense of smell more complex than ever. The postgenomic era can be defined as the period following the completion of the sequencing of the human genome, which has been dominated by “*transdisciplinarity, speed, and centrality of computational technology that mark the contemporary life science*”.[21] This means that one is able to scrutinize, with many details, the interactions between ORs and odorants using *in vivo*, *in vitro*, or *in silico* methods.

Olfactory receptors are undoubtedly the cornerstone of the specificity of the olfactory system. At the molecular level, the system is likely to obey classical pharmacological rules, where a ligand activates a given receptor with a defined potency. Multiplying these rules across several hundreds of ORs creates a subtle combinatorial code that allows for extraordinary discriminating power. Decoding this combinatorial code requires deciphering the differential activation of olfactory receptor neurons (ORNs).

Cracking the code is of course the first step towards our understanding of the complex mechanism involved in the perception of smell. Such a function is performed by the central part of our olfactory system (the olfactory bulb, olfactory cortices, etc.). Beyond that, so-called peri-receptor events, notably implying proteins present in the olfactory mucus, likely contribute to the subtlety of our perception by modulating how the quality, the quantity, and the kinetics of the odorant signal are coded.[22-24] A preliminary study based on the response of 40 ORNs shows that the quality of an odor correlates with the combinatorial code of OR activation, suggesting that the creation of structure-odor relationships would benefit from neuroscience approaches.[25] It is hoped that these

techniques will allow us to take a hypothesis-driven approach to discovering novel molecules of interest based on the properties of ORs.

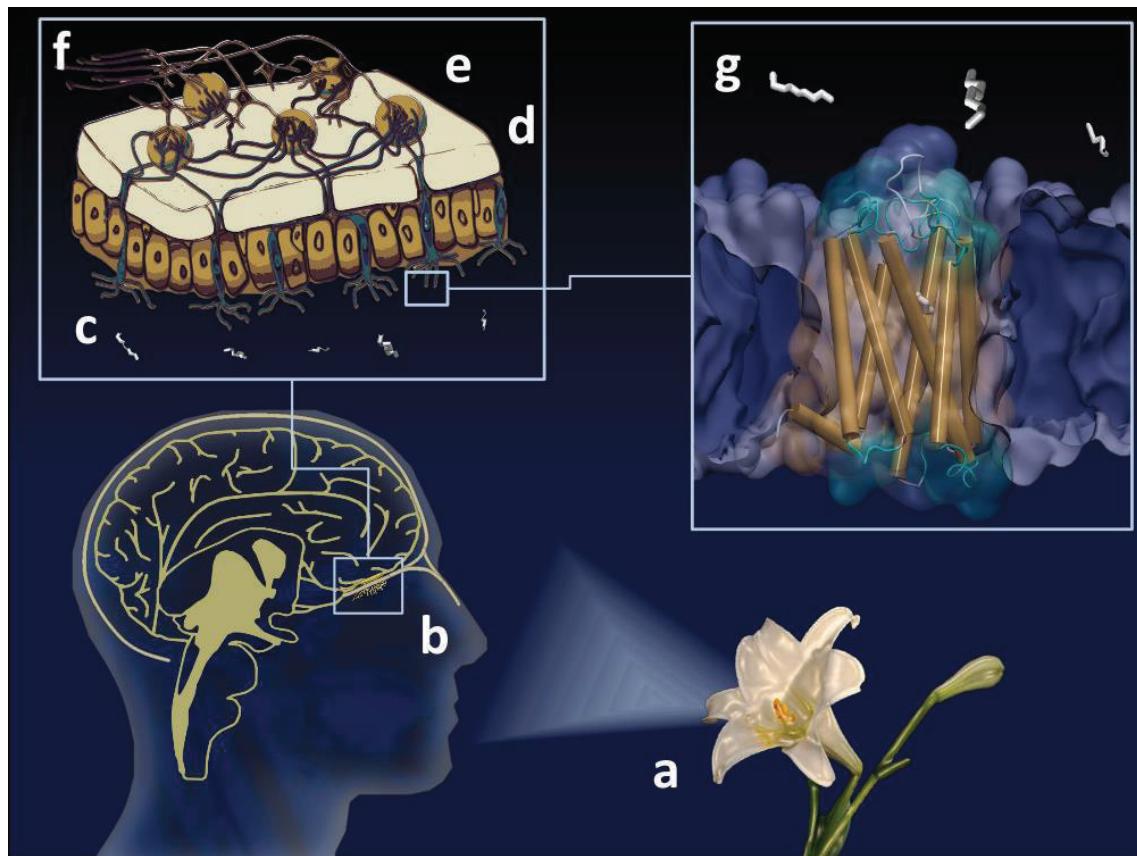


Figure 1. Process of odorant chemoreception. Emitted volatile molecules (a) are sniffed and enter in contact with our olfactory epithelium (b) located at the top of our nasal cavity. The olfactory epithelium (c) is protected from drying out by the olfactory mucus where odorant molecules will be solubilized. Within this mucus, olfactory neurons project their cilia to detect the chemical signals. Chemical signals are transformed into an electric signal, further transmitted through the axon of the neurons (d) which cross the ethmoid bone to reach the olfactory bulb. Olfactory neurons expressing the same Olfactory Receptor converge towards the same glomeruli (e). After signal processing by these glomeruli, the olfactory information spreads in various part of our brain, responsible for either conscious or emotional aspects of the perception of smell. At the molecular level, and back to the surface of the neuron, odorant molecules are involved in specific interactions with our Olfactory Receptors (g) embedded in the phospholipidic membrane of the neuron. The complementarity between the physicochemical properties of both the odorant and the Olfactory Receptor binding pocket defines the selective activation of the neuron.

In this review, we provide some clues for understanding the complexity of the recognition spectrum of olfactory receptors. We propose a survey of the recent literature regarding the choice of odorant sets used to find relationships between odorants and receptor activation. We also briefly compare *in vitro* approaches typically used to deorphanize ORs (an orphan receptor is a receptor without known

agonist), and we provide lists of already deorphanized human ORs relating to both the chemical and odorant spaces that they cover.

Interactions between ORs and odorants

OR genes are a multigenic family corresponding to more than 2% of our genome. The genes that code for our ORs were discovered by Buck and Axel in 1991.[1] In humans, their number is ~1000, of which part have been tagged as pseudogenes; to date, 396 potentially functional receptors have been identified.[19] However, our sense of smell has been associated with only 2/3 of those because 273 human olfactory receptor genes were shown to be expressed in the olfactory epithelia of 26 individuals.[26] It should be noticed that our sense of smell is thought to have become less and less important over the course of evolution, explaining the large numbers of pseudogenes in our genome. Recent studies have suggested that a plateau has not yet been reached in this decrease in functional gene number and that we will continue losing our sensing functionality over time.[27]

Based on sequence analysis, a classification has been proposed where ORs are named ‘ORN_NX_M’. ORs sharing at least 40% sequence identity belong to the same family ‘N’. Within this main family, those sharing at least 60%, belong to the same sub-family ‘X’, and ‘M’ is the number of this OR in the group.[28] For example, OR7D4 and OR1G1 share less than 40% identity, whereas OR7D4 and OR7D2 possess at least 60% of sequence identity. Interestingly, ORs are expressed in other organs [29] such as the heart,[30] male germinal cells,[31] spleen, pancreas,[32] blood leukocytes,[33] and kidney.[34] Reviews of these ectopic expressions can be found in ref [35, 36]. In those cases, the term ‘odorant receptor’ is semantically more accurate than ‘olfactory receptor,’ as the first term only emphasizes the chemo-genomic link, i.e., this family of receptors is associated to chemicals belonging to the family of *odorant* molecules. In contrast, ‘olfactory’ suggests that these receptors would be specifically expressed within the olfactory system.

Odorant sets

The physicochemical criteria required for a chemical to belong to a family of odorants are rather wide: a molecular weight under ~400 g.mol⁻¹ and a weak polarity associated with a high lipophilicity while maintaining a certain water solubility.[37] As a consequence, the odorant chemical space is virtually infinite due to the large number of chemical groups that fulfill these criteria. False-positives can nevertheless be found, with chemicals that fulfill these criteria but do not trigger any smell, such as propane. From a biochemical point of view, these properties can be connected to their relationship with ORs. The odorant should first be able to reach our nasal cavity through the inspired air (low molecular weight), solubilized within the olfactory mucus (solubility), and finally and most importantly, be the agonist of an OR with a hydrophobic cavity (lipophilicity).

When trying to deorphanize ORs, it is difficult to work from a set of compounds that covers all chemical facets of odorant molecules. Based on chemometric analyses, the concept of diversity has been proposed for sets of odorants,[38] which describes the chemical space covered by odorants and ensures that the set will not miss any chemical features that could activate a receptor. This strategy was used to screen many receptors while trying (with limited success) to identify the odorant space associated with some OR sequences.[39] It appears that the chemical space of odorants is quite subtle and cannot be fitted by simple mathematical laws, even when elaborate protocols are used. As such, it is highly unlikely that prototypical odorant can be identified that would represent a chemical family for all odorant receptors. Although hexanal can be regarded as prototypical of medium size aliphatic aldehydes (heptanal, octanal) for OR1G1,[40] OR1A1 and OR1A2 respond completely differently within this series.[41] Using physicochemical descriptors, Mainland et al. selected 73 odorants chosen to represent a set of 2728 chemicals for use in screening 394 human ORs, and fewer than 7% were associated with agonists, revealing the extreme difficulty in simplifying the odorant space.[42]

Using a different strategy, odorants were placed into several categories that would be of interest to the flavor and fragrance communities rather than seeking a complete chemical sampling. Krautwurst *et al.* ranked a set of 285 odorants into those of potential interest for flavor science (121 called key food odorants, KFOs), those related to body odors (28 BO) or others (158 molecules).[43] They were further shown to be the best candidates for OR deorphanization due to their capacity to trigger their activations. This was calculated two times in comprehensive meta-analyses, leading to the unambiguous result that cognate key food odorant-OR pairs are about three times more frequent than those involving non-key food odorants.[43, 44] Authors put forward the intriguing relationship between the number of key food odorant (230 KFOs) and that of functional ORs genes expressed in humans (273 ORs, *vide infra*). We are not convinced that the number of key food odorant has to be so directly related to the number of functional OR genes within a species. Typically this would suggest that dogs, rats and even elephants would have a larger foodborne stimulus space than us, as they express around 800, 1200 and 2000 functional OR genes, respectively.[45]

Cellular expression systems for the study of ORs

The detection of OR activation in a physiologically relevant medium such as a neuron or directly within the brain is very challenging. As a result, various *in vitro* assays have been developed as useful tools to isolate and quantify the response of a receptor following its stimulation with odorants. Despite difficulties in expressing ORs within heterologous cells, several approaches have been successful using genetically modified expression systems that facilitate the recording of the functional activities of a number of ORs.

Olfactory receptor neurons (ORNs) themselves would intuitively be the most effective expression system for various ORs.[46] The use of cell lines derived from ORN progenitors has also been described, and this method appears to be more suitable at a reasonable cost.[47] However, the heterogeneity of the OR population in a primary ORN culture system and difficulties in their maintenance limit their use. The OR coding sequence that is used for the expression of a single OR in each ORN suppresses the expression of multiple ORs, which could affect their heterogeneous expression.[48]

Human embryonic kidney (HEK) 293 cells are good candidates for the heterologous expression of ORs, and they show high efficiency for growth and transfection.[49] Indeed, they are generally used for the functional expression of various types of G protein-coupled receptors (GPCRs). Nevertheless, the stable trafficking of ORs to the plasma membrane for cell surface expression is the most challenging task in studying the functional activities of ORs. A HEK293-derived cell line named Hana3A expressing ORs has been shown to perform well in this task.[50] Hana3A stably expresses receptor transporting proteins 1 and 2 (RTP1 and RTP2) and receptor expression enhancing protein 1 (REEP1)[51] to facilitate the trafficking of ORs to the plasma membrane. These cell lines also stably express the homologous G protein that couples to ORs (G_{olf} α subunit, $G_{\alpha\text{-olf}}$), which avoids artifacts that can result from the coupling of the OR with heterologous G protein subunits, which could modify the response of the cell upon odorant stimulation.[52, 53]

Alternatively, Xenopus oocytes are widely used for expressing membrane receptor proteins and channels because they do not express many endogenous ion channels or receptors, thus avoiding perturbations. Electrophysiological recordings have been used with the Xenopus oocyte system for the purpose of pharmacological analyses, biophysical investigations, and receptor mutagenesis studies.[54, 55] Several studies have already demonstrated that ORs from various species can be expressed in oocytes for the purpose of analyzing their ligand specificity.[56-61] Another advantage to using the oocyte expression system is its feasibility for studying the functional interaction of ORs with downstream components via co-injection approach into a single-cell.

The use of Sf9 cells, which are non-mammalian cells originating from moths, is another way to express ORs.[62] hOR1G1 was expressed in Sf9 cells to measure their calcium discharge in response to various chemical stimulations.[63-65] Despite this series of success in heterogeneous OR expression, it should be stressed that the heterogeneous expression system or the nature of the G protein leads to variability in the list of measured agonists for a given OR (see the review by Peterlin et al.).[66]

The yeast functional expression system has been demonstrated as being feasible for studying ORs due to its null background for membrane receptors.[67-72] Yeast also provides a sensitive reporter

system with the functional homologies between yeast pheromone and mammalian GPCR signaling.[73] High-throughput screening of ORs has been performed using the yeast *Saccharomyces cerevisiae* as a host system.[74]

As discussed in a review focused on OR deorphanization, all of these *in vitro* methods have produced different results, either due to the expression system used, the coupling with the G protein or the readout to measure the OR response.[66] Very recently, an *in vivo* deorphanization was reported ('Kentucky *in vivo* odorant ligand-receptor assay') that could prevent such artifacts.[11]

Interestingly, *in vitro* expression has been used to identify the unknown function of certain ORs expressed in non-olfactory tissues. In the case of spermatogenic cells (expressing OR1D2), Spehr *et al.* expressed the receptor in HEK293 cells,[75] and they determined that sperm chemotaxis is controlled by hOR1D2 by examining the behavior of sperm swimming upon bourgeonal stimulation. The biological agents involved in the chemotaxis of sperm have been the subject of numerous studies because the latency of the OR response did not correspond with the velocity of the effective Ca^{2+} influx. In 2011, a study showed that the chemotaxis of sperm was controlled by the CatSper Ca^{2+} channel activated by progesterone, a female hormone.[76] Brenker *et al.* published that, in addition to progesterone, odorants are also able to directly activate the crucial CatSper Ca^{2+} channel, reconciling both studies.[77]

Receptor pharmacology

ORs belong to the family of class-A G-protein coupled receptors (GPCRs), which are seven transmembrane proteins that detect the presence of ligands at the surface of neurons. As with all GPCRs, the function of ORs can be rationalized from a pharmacological point of view.[78]

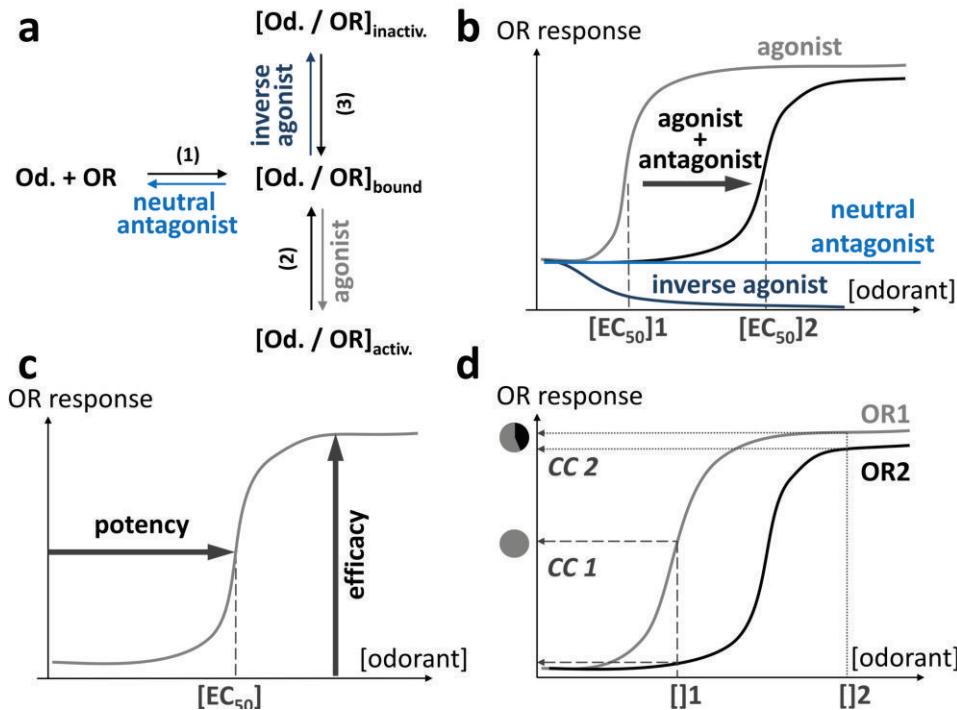


Figure 2. Pharmacology of Olfactory Receptors. (a) Thermodynamic equilibria associated to odorant categories. An odorant and an OR are in a chemical equilibrium (1) that defines if the odorant is a neutral antagonist or not. If the odorant is bound to the OR, the system will be subjected to two equilibria. If the ligand favors the active state (2), it is an agonist. If the odorant favors the inactive state of the receptor, it is an inverse agonist (3). These equilibria can eventually be modified by allosteric modulators (acting out of the binding site), such as Odorant Binding Proteins (see text). (b) The mixture between an agonist and one of its antagonists will modify the response of the receptor upon stimulation. The antagonist will shift the potency (EC₅₀, see (c)) of the agonist towards that of a weaker ligand. With such effects within the combinatorial code of OR activation, the perception of the mixture can be different from that of the sum of isolated compounds. An inverse agonist will decrease the spontaneous activity of the receptor, while the neutral antagonist won't perturb the receptor. (c) Definitions of potency and efficacy. The potency is connected to the [EC₅₀] and is defined by the concentration of odorant necessary to a half-activation of OR. The maximum level of OR activation for a given ligand is called efficacy. (d) Effect of concentration of an odorant on typical dose-response curves of two ORs. In this example, at low concentration ([]1) an odorant triggers only the activation of OR1, at higher concentration ([]2) OR2 is also responding and now also contributes to the combinatorial code (CC) associated to the perception.

When a ligand potentially interacts with a receptor, the system is subjected to several chemical equilibria that will define the nature of the ligand (Figure 2a). Odorants can be split into several categories for a given OR, *viz.*: agonists, neutral antagonists, non-binders, and inverse agonists (Figure 2b). While non-binders do not exhibit any affinity for the receptor (meaning that they do not trigger any response), agonists trigger a response of the receptor proportional to their potency and efficacy (Figure 2c). Antagonists have the opposite effect, as they compete with agonists within the

receptor. The last category that can also be considered as odorant modulators is that of inverse agonists, which decrease the basal activity of a receptor (Figure 2b). Neutral antagonists do not change the OR response upon stimulation. In GPCRs, positive or negative allosteric modulators can also be found. In those cases, the ligand does not bind within the canonical binding cavity of the receptor but triggers an increase or a decrease in the response of the receptor through an alternative binding site. A typical example was reported through the modulation of an OBP on the shape of the dose-response curve between a ligand and an OR.[79] From a general manner, the role of OBP is a good example of OR response modulation.[79-82] Non-competitive antagonism between odorants has also been reported to act through an alternative signalization pathway within the cell.[83] These pharmacological features might explain the otherwise confusing perception modulations that have been observed.

The very low detection threshold of sulfur compounds could be partly due to copper-mediated ORs, as demonstrated in the mouse OR244-3;[84, 85] in humans, the orthologous receptor is OR4E2. *In vitro* experiments on hOR4E2 have reported that it responds to many types of ligands,[42] but to date, no data on sulfur compounds have been obtained to confirm that it is copper-dependent and that it responds to sulfur compounds.

The evolution of an odor based upon the odorant concentration or its association to a mixture can be rationalized by examining its potency and efficacy. Concerning concentration-dependent perception, the example of the ‘cat ketone’ is typical. This molecule (4-mercaptop-4-methylpentan-2-one) can have a bad smell like cat urine at high concentrations,[86] but the perceived odor becomes ‘cabernet-sauvignon’ or blackcurrant when the concentration is low.[87] This change in perception suggests that the combinatorial code of OR activation differs upon concentration (Figure 2d). In fact, examples of such a differential perception upon variations in the concentration of the odorant have been known for years.[88] Note that this dose-dependent response of ORs and their patterns of activation can be projected to those of glomeruli.[89-91]

In the case of hOR5P3, *in vitro* experiments have identified more molecules inhibiting OR signaling than molecules triggering OR activation. This suggests that this receptor mainly contributes to the combinatorial code of the perception of odors through negative responses upon stimulation (*vide infra*).[92] This kind of response can either be interpreted as an inverse agonist behavior or a receptor non-specific signaling inhibition. Adapted experimental controls should be performed to demonstrate the true inverse agonist nature of the molecule. In some cases however, an OR does not have a strong basal activity, suggesting that the concept of inverse agonist cannot be generalized to all ORs. These molecules can eventually compete with agonists, as shown for cycloheptane-carbaldehyde (chca), which decreases the basal activity of the rat ORI7 and also strongly decreases

the response of the receptor to a strong agonist. For unknown reasons, this modulation is only observed when the inverse agonist is applied prior to stimulation by the agonist.[93]

The identification of ligand effects (agonist, antagonist, non-agonist or inverse agonist) on a receptor is generally achieved through *in vitro* experiments. Unfortunately and as previously reported, these effects can differ depending on the experimental protocol. For example, nonanal has been reported as both an agonist and an inverse agonist for hOR1A1, suggesting that the combinatorial code of a pure compound must be determined with caution.[41, 92] Further investigations are needed to assess which *in vitro* protocol allows for the greatest accuracy in capturing *in vivo* pharmacology.[66]

Of course, the effects produced by mixtures of compounds are much closer to what occurs in real life. This situation is also more complicated, involving interactions between messages from receptors within the central neuronal networks of the olfactory system. Nevertheless, some odorants can act as agonists for some ORs and as antagonists for others.[40, 94] Again, the mixture of two odorants is likely to trigger a smell that is different from the sum of the two independent chemicals due to a modification of the combinatorial code through antagonistic effects. For example, such effects have been reported for the human OR1G1 receptor with mixtures of whiskey lactone and isoamyl-acetate or the mouse eugenol receptor in response to isoeugenol and methyl-isoeugenol.[40, 94-96] At the moment, no rules exist to predict the agonist/antagonist/inverse agonist actions for a receptor, and these effects emphasize the virtual impossibility to deconstruct the sense of smell into a simple sum of independent stimulations. Basic research is still needed to understand more deeply the mechanism of GPCR activation upon ligand stimulation.

Chemo-genomic links

Focusing on flavors and fragrances, connections between our perception of chemicals and our genome have been tentatively established. However, the relationships between OR activation and the evoked smell are hard to establish. It is likely that the contributions of some ORs are more important than others, as we will discuss below. Functional OR genes can indeed vary between humans, possibly due to a lack of natural evolutionary pressure. In several cases, variations in sequences do alter the *in vitro* response to some odorants, as has been shown by studying the *in vitro* function of 18 different ORs.[42] Most of the time, the impact of these variations in genotype on the OR phenotype is minor, and the selectivity of a receptor and its mutants is conserved. Rather, the primary difference is in the activation threshold, as is the case for cis-3-hexenol, the perception of which is impacted by polymorphisms in OR2J3.[97] The perception of androstenone is more well documented, even if a significant population is nearly anosmic to this compound.[98] Although the majority of people who detect androstenone describes it as animal and/or urine, some people expressing a variant in OR7D4 that differs by two amino-acids (on a sequence of ~300 residues)

reports the smell to be more pleasant, with vanilla and honey notes. Other people with a single mutation within the sequence can be considered to be “super-smellers,” with a detection threshold lower than average.[99] In the case of this receptor, this modulation in perception is associated with a preference for meat coming from pork, whether castrated (with a low amount of androstenone) or not (with a higher amount of androstenone).[100] Interestingly, the genotype associated with this receptor results in a phenotype that itself triggers anthropological behaviors. Indeed, a correlation can be observed between the ratio of people who are anosmic to androstenone in France, Spain and United Kingdom and the percentage of castrated pork in those countries.[101] It has also been shown that this modulation in perception is a heritable trait[102] and is a prime example of how a single OR can be dominant within the combinatorial code of an odorant.

Similarly, the differential sensitivity to isovaleric acid amongst the population could be partly due to a polymorphism within the gene expressing OR11H7P.[61] Extending such studies to the perception of foods and beverages, the odor of β -ionone, both pure and within food, is affected by polymorphisms within the OR5A1 gene.[103] Perfectly illustrating the role of genomics within the field, a statistics analysis was performed on ~15,000 individual genomes of individuals whose hedonic perception of coriander (also known as cilantro) was known. The dislike of coriander is related to a sequence variation within the OR6A2 gene, which impacts the sensitivity to aldehydes present in coriander leaves. However, the heritability of this perception has not been established.[104] Less specifically, the perception of the typical smell of urine after having ingested asparagus, driven by methanethiol and associated odorant molecules, was partly associated with variations in OR7M2 and less so in OR14C36 in Caucasian people, whereas in African people, no association could be established.[105] In general, the ability to associate odorant molecules to certain ORs would greatly aid in deciphering the preferences in flavors or fragrances amongst the population.

Computational approaches

The virtually infinite odorant space makes the experimental testing of all odorants for any OR impractical. Modeling approaches can be adapted to guide the screening of OR/odorant associations. Consequently, atomic-level approaches have been tentatively developed to predict the phenotypes associated with an OR when it is stimulated by a candidate ligand. Unfortunately, no experimental structure of any OR is available to date. The only rigorous information about ORs is their sequences, which has already allowed us to gain insights into their putative binding sites for the purpose of predicting their associated chemical spaces.[106] For several years, 3D models have become more accurate as they have been built under the constraint of experimental *in vitro* data. Over the last few years, models of ORs have provided accurate atomic-level details for their responses to chemically related odorants via modifications in the carbon skeleton,[107] by odorants with nitro groups[108] or

by odorants belonging to different chemical classes[64]. Models have also helped to propose the reprogramming of an OR for a molecule that was initially a non-agonist.[109] Only recently have *in silico* approaches been used for the purpose of OR deorphanization. A detailed review on modeling approaches has been published recently.[110] As with all techniques, a trade-off between speed and accuracy must be made. By using a rapid screening based on docking (which was already used for the purpose of mORs deorphanization)[111] on a large database of more than 500 odorants, researchers were able to predict 40 potential mOR42-3 agonists of which half were experimentally assessed.[112] Although not perfect, this study suggests that *in vitro*-guided experiments will certainly help in the search for the odorant space associated with a receptor. Using a more elaborate but also more time consuming method based on statistical thermodynamics, an accurate sorting of a series of 10 ligands into 8 agonists and 2 non-agonists for a broadly tuned human receptor was realized.[65] Generally, these atomic-level models can describe how very subtle chemical differences between odorants can be discriminated by ORs, which explains the difficulties of approaches aimed at proposing universal models of structure-odor relationships by only focusing on the odorant structure. In a simple model, an odorant can be regarded as a vector made up of 396 (if allelic variation is omitted) dimensions corresponding to each human OR response. The paradigm of the perception of smell would be that similar vectors are associated with similar odors. As cited earlier, encouraging signs have been experimentally obtained on a limited set of ORs and odorants[25], but this hypothesis can be validated only when such experimental or sufficiently accurate theoretical data are available. Very few studies have provided structure-function of mammalian ORs where *in silico* models have been supported by site-directed mutagenesis experiments, showing that modeling protocols were already sufficiently accurate.[41, 85, 96, 109, 113, 114] It is likely that with the increase of computational power, the availability of experimental templates and more accurate energy prediction protocols, modeling studies will provide accurate clues that will be useful in the design of agonists, antagonists or inverse agonists for ORs.

OR deorphanization

Using *in vitro* recordings of the activation of ORs and sometimes with only one discovered ligand, 57 human ORs have been associated to chemicals thus far. This represents slightly less than 15% of the unaltered 396 human OR genes. The different methodologies used for the *in vitro* assays makes their comparison tricky, as the chemical space associated with an OR can vary with the expression system used or the G-protein coupling.[66] In addition, the difficult task of identifying an odor descriptor for a chemical was done with the purpose of identifying connections between OR activation and the odor space. In Table 3, we report the names of agonists along with their primary odor descriptor. Most descriptors associated with each chemical in the table are taken from the literature, and the

description was determined based on GC-olfactometry studies. In a few cases where GC-olfactometry was not available, the note is taken from the description of an expert. We chose to select only one term from the full list of descriptors, checking that the term is sufficiently general. The descriptors are further categorized within the following list: floral, fruity, animal, sulfur, green, balsamic, spicy, and woody, which are mainly used by flavorists and perfumers.[115] The chemicals used for the purpose of OR deorphanization were not systematically purified, and the contamination by unwanted odorants coming from either the experimentalist or from trace amounts within the test set cannot be excluded. The derivatives of an odorant can contribute to modifying the response of a receptor *in vitro*, as shown from stored vs. freshly purified eugenol against mouse OR-EG.[94] Mixtures of enantiomers can also be mistakenly considered to be pure compounds, and the information about the enantiomeric mixture is generally not specified. Notice that in many cases validation of OR-odorant interaction is not performed through dose-response analysis.

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

Receptor Name	agonist	Ref agonist	Ref odor
OR10A6	3-phenyl propionate	[42]	floral -
	androstadienone	[42]	musky, urine [99]
OR10G3	cinnamaldehyde	[42]	cinnamon [116]
	ethyl vanillin	[42, 92]	vanilla [117]
OR10G4	eugenol	[92]	spicy [116]
	vanillin	[42]	vanilla [118]
OR10G7	ethyl vanillin	[92]	vanilla [117]
	vanillin	[42]	vanilla [118]
OR10G9	ethyl vanillin	[92]	vanilla [117]
OR10J5	lyral	[39, 42, 92]	floral [121]
OR11A1	2-ethyl fenchol	[42, 92]	earthy [122]
OR11H4	isovaleric acid	[61]	cheese [123]
OR11H6	isovaleric acid	[61]	cheese [123]
OR11H7P	isovaleric acid	[61]	cheese [123]
OR1A1	(-)carveol	[41]	spicy [124]
	(-)carvone	[39]	green [125]
	(+)-carvone	[39, 42, 92]	mint [126]
	(+)-dihydrocarvone	[39]	herbal [127]
	(+)-menthol	[42]	mint [128]
	(R)-(+)citronellol	[41]	floral [129]
	(S)-(+)citronellal	[41]	citrus [130]
	(S)-(+)citronellol	[39, 41]	rose [131]
	1-decanol	[39]	polished [131]
	1-heptanol	[39]	green [131]
	2-heptanone	[39]	banana [125]
	2-nonenone	[39, 42]	green [132]
	2-octanone	[39]	mushroom [124]
	3-heptanone	[39]	roasty [133]
	3-octanone	[39]	mushroom [123]
OR1C1	3-phenyl propyl propionate	[42]	floral -
	4-chromanone	[39]	- -
	4-decenal	[41]	fried [134]
	allyl heptanoate	[39]	fruity [135]
	allyl phenylacetate	[39, 42, 92]	honey -
	androstadienone	[42]	musky, urine [99]
	benzophenone	[39]	floral [136]
	benzyl acetate	[39]	banana [133]

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

Receptor Name	agonist	Ref agonist	Ref odor
OR1A2	beta-damascone	[42]	rose [137]
	bourgeonal	[138]	floral [107]
	butyl anthranilate	[42]	fruity [139]
	cinnamaldehyde	[42]	cinnamon [116]
	citral	[41]	lemon [140]
	coffee difuran	[42]	coffee [141]
	dihydrojasmine	[39]	floral [142]
	ethylene brassylate	[42]	musky [143]
	eugenol	[42]	spicy [116]
	geraniol	[39, 41, 138]	rose [144]
	geranyl acetate	[42]	rose [145]
	helional	[41, 92, 138]	floral [146]
	heptanal	[41]	fatty [123]
	hydroxy-citronellal	[41]	floral [147]
	limonene	[42, 92]	lemon [120]
	methyl furfuryl disulfide	[42]	alliaceous [148]
	nonanal	[41]	orange [149]
	nonanethiol	[39]	stinky [150]
	octanal	[41]	watermelon [131]
	octanethiol	[42]	stinky [150]
	octanol	[41]	soap [131]
	quinoline	[42, 92]	coal tar [148]
	shoyu pyrazine	[42]	earthy [151]
	terpineol	[42]	pine [131]
	terpinyl acetate	[42]	waxy [128]
	(-)-carveol	[41]	spicy [124]
	(R)-(+)citronellol	[41, 152]	floral [129]
	(S)-(+)citronellal	[41, 152]	citrus [130]
	4-decenal	[41]	fried [134]
	citral	[41, 152]	lemon [140]
	geraniol	[41]	rose [144]
	helional	[41]	floral [146]
	heptanal	[41]	fatty [123]
	hydroxy-citronellal	[41]	floral [147]
	nonanal	[41]	orange [149]
	octanal	[41]	watermelon [131]
	octanol	[41]	soap [131]
	linalool	[42, 92]	bergamot [124]
	androstenone	[92]	musky, urine [99]
	coumarin	[92]	woodruff [120]
	nonanoic acid	[92]	coriander [153]
	(4-tert-butylphenoxy) acetaldehyde	[75]	- -
OR1D2	3-phenylbutyraldehyde	[75]	green [154]
	3-phenylpropionic aldehyde	[75]	green [155]
	4-phenylbutyraldehyde	[75]	rose [154]

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
	de			
	bourgeonal	[75, 138, 156]	floral	[107]
	canthoxal	[75]	licorice	[115]
	cyclamal	[75]	floral	[115]
	floral ozone	[75, 156]	ocean	[115]
	lilial	[75, 156]	floral	[121]
	phenyl acetaldehyde	[75]	lilac	[118]
OR1E3	acetophenone	[63, 138]	vegetal	[157]
	(+/-)-citronellal	[40]	citrus	[130]
	1-decanol	[40]	polished	[131]
	1-dodecanol	[40]	raw carrot	[157]
	1-heptanol	[40]	green	[131]
	1-hexanol	[40]	green	[131]
	1-nonanol	[40, 64]	green	[158]
	2-ethyl-1-hexanol	[40]	rose	[133]
	2-isobutyl-3-methoxypyrazine	[40]	green	[159]
	2-methyl pyrazine	[40]	roasted nuts	[157]
	2-nonenol	[40]	plastic	[160]
	2-nonenone	[40]	green	[132]
	2-octanol	[40]	fatty	[161]
	2-undecanone	[65]	floral	[158]
	3-hydroxybutan-2-one	[40]	buttery	[123]
	3-nonenone	[40]	green	[162]
	3-octanol	[40]	earthy	[143]
	4-octanol	[40]	soapy	[163]
	9-decen-1-ol	[64]	waxy	[164]
OR1G1	acetophenone	[40, 138]	vegetal	[157]
	benzaldehyde	[40]	almond	[131]
	benzothiazol	[40]	rubbery	[157]
	beta-ionone	[138]	violet	[120]
	camphor	[40, 64]	camphor	[158]
	capric acid	[40]	waxy	[165]
	cinnamaldehyde	[166]	cinnamon	[116]
	citral	[40]	lemon	[140]
	coumarin	[40]	woodruff	[120]
	decanal	[40]	waxy	[133]
	ethyl butyrate	[40]	strawberry	[131]
	ethyl decanoate	[40]	grape	[167]
	ethyl isobutyrate	[40, 65]	fruity	[168]
	ethyl nonanoate	[40]	waxy	[143]
	ethyl octanoate	[40]	waxy	[169]
	ethyl vanillin	[40]	vanilla	[117]
	eugenyl acetate	[166]	spicy	[119]
	floral ozone	[166]	ocean	[115]
	gamma-decalactone	[40]	peach	[131]
	geraniol	[40, 138]	rose	[144]
	guaiacol	[40]	smoky	[149]
	hedione	[40]	floral	[170]
	heptanal	[40]	fatty	[123]

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
	hexanal	[40]	green apple	[123]
	isoamyl acetate	[40, 65, 138]	overripe banana	[171]
	jasmonyl	[166]	jasmine	[115]
	lauric aldehyde	[40]	waxy	[133]
	limonene	[40]	lemon	[120]
	lyral	[40]	floral	[121]
	maltol	[40]	caramel	[172]
	maltyl isobutyrate	[166]	fruity	[115]
	manzanate	[166]	apple	[115]
	menthol	[40, 63]	mint	[128]
	methyl decanoate	[40]	floral	[173]
	methyl nonanoate	[40]	fruity	[148]
	methyl octanoate	[40]	green	[159]
	nonanal	[40, 64]	orange	[123]
	nonanoic acid	[40]	coriander	[153]
	octanal	[40]	watermel- on	[131]
	octanol	[40]	soap	[131]
	phenylmethanol	[40]	green	[174]
	piperonyl acetone	[40]	floral	[175]
	pyrazine	[40]	nutty	[176]
	pyridin	[40]	scallop	[132]
	quinoline	[40]	coal tar	[148]
	safrole	[40]	spicy	[177]
	S-methylthio butanoate	[40]	cheesy	[178]
	thiazol	[40]	meaty	[179]
	thymol	[40, 138]	spicy	[157]
	trans-anethol	[40]	anise	[180]
	tridecanal	[65]	perfum- like	[126]
	vanillin	[40]	vanilla	[118]
OR1L3	alpha-damascone	[166]	fruity	[139]
	vanilin	[166]	vanilla	[118]
OR2A25	androstadienone	[42]	musky, urine	[99]
	geranyl acetate	[42, 92]	rose	[145]
	quinoline	[92]	coal tar	[148]
OR2AG1	amylbutyrate	[181, 182]	fruity	[167]
OR2AT4	brahmanol	[183]	woody	[115]
	sandalore	[183]	woody	[115]
	1,3-butane dithiol	[42]	alliaceous	[184]
OR2B11	4-hydroxycoumarin	[92]	woodruff	[120]
	cinnamaldehyde	[42]	cinnamon	[116]
	coffee difuran	[42]	coffee	[141]
	coumarin	[42, 92]	woodruff	[120]
	limonene	[92]	lemon	[120]
	quinoline	[42, 92]	coal tar	[148]
OR2B3	beta-ionone	[166]	violet	[120]
	eugenyl acetate	[166]	spicy	[119]
	nerolidol	[166]	floral	[139]
OR2C1	nonanethiol	[39]	stinky	[150]
	octanethiol	[39, 42]	stinky	[150]

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
OR2G2 	alpha-damascone [166] cinnamaldehyde [166] maltyl isobutyrate [166] vanilin [166]		fruity [139] cinnamon [116] fruity [115] vanilla [118]	
	(+)-carvone [42] (+)-menthol [92] 1-decanol [39] 1-heptanol [39] 1-nonanol [39] 2,4-DNT [42, 92] 2-methoxy pyrazine [42] 2-nonenone [42] 3-phenyl propyl propionate [42] 4-hydroxycoumarin [39]		mint [126] mint [126] polished [131] green [131] green [158] - nutty [185] green [132] floral - woodruff [120]	
OR2J2 	androstadienone [42] butyl anthranilate [42] capric acid [42] cinnamaldehyde [42] citral [92] coffee difuran [42] coumarin [42, 92] ethyl vanillin [42, 92] eugenol [42] eugenol methyl [92] eugenol acetate [92] geranyl acetate [42, 92] helional [92] nonanal [92] octanol [39, 92] octanethiol [92] phenyl acetaldehyde [42] quinoline [42, 92] vanillin [42]		musky, urine [99] fruity [139] waxy [165] cinnamon [116] lemon [125] coffee [141] woodruff [120] vanilla [117] spicy [116] smoky [120] spicy [119] rose [145] floral [146] orange [123] soap [131] stinky [150] lilac [118] coal tar [148] vanilla [118]	
OR2J3 	2,4-DNT [42, 92] 2-nonenone [42] 3-phenyl propyl propionate [42] cinnamaldehyde [42] cis-3-hexenol [97] citral [92] coffee difuran [42] coumarin [42, 92] ethylene brassylate [42] eugenol methyl [92] geranyl acetate [42, 92] helional [92] octanol [92] quinoline [42]		- green [132] floral - cinnamon [116] green [171] lemon [125] coffee [141] woodruff [120] musk [143] spicy [120] rose [145] floral [146] soap [131] coal tar [148]	
OR2M2 	(S)-(-)-citronellol [39]		rose [131]	
OR2M4 	alpha-damascone [166] cinnamaldehyde [166] cresyl methyl ether [166]		fruity [139] cinnamon [116] naphthyl [139]	

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
OR2M7 	estragole [166] fructone [166] nerolidol [166] vanilin [166]		anise [139] fruity [115] floral [139] vanilla [118]	
OR2T10 	(S)(-)-citronellol [39] geraniol [39] alpha-damascone [166] cinnamaldehyde [166] maltyl isobutyrate [166] terpinyl acetate [166] vanilin [166]		rose [131] rose [144] fruity [139] cinnamon [116] fruity [115] waxy [128] vanilla [118]	
OR2T34 	alpha-damascone [166] cinnamaldehyde [166] estrugole [166] floral ozone [166] fructone [166] jasmonyl [166] vanilin [166]		fruity [139] cinnamon [116] anise [139] ocean [115] fruity [115] jasmine [115] vanilla [118]	
OR2W1 	(-)carvone [39] (+)-carvone [39, 42, 92] (+)-dihydrocarvone [39] (+)-menthol [42] (S)(-)-citronellol [39] 1,1-dimethoxy-octane [42] 1,3-butane dithiol [42] 1-decanol [39] 1-heptanol [39] 1-hexanol [39] 1-nonanol [39] 2,3-hexanedione [39] 2,4-DNT [42] 2-ethoxythiazole [42] 2-heptanone [39] 2-hexanone [39] 2-methoxy pyrazine [42] 2-nonenone [39, 42] 2-octanone [39] 3,4-hexanedione [39] 3-heptanone [39] 3-octanone [39] 3-phenyl propyl propionate [42] 4-chromanone [39] 4-hydroxycoumarin [39] acetophenone [39] allyl phenylacetate [39, 42, 92] amyl acetate [42] benzophenone [39] benzyl acetate [39] butyl anthranilate [42] butyl butyryl [42]		green [125] mint [126] musty [127] mint [126] rose [131] green - alliaceous [184] polished [131] green [131] green [131] green [158] buttery [139] - vegetable [143] banana [125] ethereal [186] nutty [185] green [132] mushroom [124] buttery [187] roasty [133] mushroom [123] floral - - woodruff [120] vegetal [157] honey - fruity [188] floral [136] banana [133] fruity [139] buttery [189]	

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
	lactate			
	butyl formate	[39]	ethereal	[190]
	capric acid	[39, 42]	waxy	[165]
	caprylic acid	[39]	fatty	[165]
	cinnamaldehyde	[42]	cinnamon	[131]
	cis-3-hexenol	[42]	green	[116]
	coffee difuran	[42, 92]	coffee	[141]
	coumarin	[42, 92]	woodruff	[120]
	decanal	[42]	waxy	[133]
	dihydrojasnone	[39]	floral	[142]
	ethyl vanillin	[42]	vanilla	[117]
	eugenol	[42]	spicy	[116]
	eugenol methyl	[92]	smoky	[120]
	geraniol	[39]	rose	[144]
	geranyl acetate	[42]	rose	[145]
	helional	[92]	floral	[146]
	heptanal	[39]	fatty	[123]
	hexanal	[39]	green	[123]
	hexyl acetate	[39]	apple	
	isovaleric acid	[42]	banana	[171]
	limonene	[42, 92]	cheese	[123]
	methyl furfuryl disulfide	[42]	lemon	[157]
	nonanal	[39]	alliaceous	[148]
	nonanethiol	[39]	orange	[123]
	nonanoic acid	[39, 92]	stinky	[150]
	octanal	[39]	coriander	[153]
	octanethiol	[39, 42, 92]	watermelon	[131]
	octanol	[39, 92]	stinky	[150]
	octyl octanoate	[42]	soap	[131]
	prenyl acetate	[39]	fruity	-
	quinoline	[42]	fruity	[191]
	terpineol	[42]	coal tar	[148]
	terpinyl acetate	[42]	pine	[131]
	aldehyde TPM	[192]	waxy	[128]
	cyclosal	[192]	fruity	[115]
	foliaver	[192]	floral	[115]
	helional	[57, 138, 192]	anise	[115]
	lilial	[192]	floral	[146]
OR3A1	methyl-hydrocinnamaldehyde	[192]	floral	-
	methyl-phenyl-pentanal	[192]	trifernal	[115]
OR4D6	beta-ionone	[103]	violet	[120]
OR4D9	beta-ionone	[103]	violet	[120]
OR4E2	amyl acetate	[42]	fruity	[188]
OR4Q3	butyl anthranilate	[42]	fruity	[139]
	coffee difuran	[42]	coffee	[141]
	eugenol	[42]	spicy	[116]
	geranyl acetate	[42]	rose	[145]

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
OR51E1	(+)-menthol	[92]	minth	[126]
	2,4-DNT	[92]	-	-
	butyl lactate	[39]	buttery	[189]
	3-methyl-valeric acid	[193]	animal	[139]
	4-methyl-valeric acid	[193]	fruity	[190]
	butyric acid	[92]	cheesy	[123]
	DMDS	[92]	sulfur	[125]
	eugenol methyl	[92]	smoky	[120]
	eugenyl acetate	[92]	spicy	[119]
	isovaleric acid	[42]	cheese	[123]
	methyl furfuryl disulfide	[92]	sulfur	[148]
	methyl salicylate	[92]	wintergreen	[189]
	nonanoic acid	[39]	coriander	[153]
	pentanol	[92]	green	[131]
	propanal	[92]	pungent	[148]
	pyrazine	[92]	nutty	[176]
OR51E2	propionic acid	[39]	vinegar	[194]
	allyl phenylacetate	[39, 42, 92]	honey	-
OR51L1	androstadienone	[42]	musky, urine	[99]
	caproic acid	[39]	goaty	[195]
	phenyl acetaldehyde	[42]	lilac	[118]
	1-nonanol	[40]	green	[158]
	2-decanone	[40]	fruity	[196]
	2-isobutyl-3-methoxypyrazine	[40]	green	[159]
	2-nonanol	[40]	plastic	[160]
	3-hydroxybutan-2-one	[40]	buttery	[123]
	3-nonanone	[40]	green	[162]
	3-octanone	[40]	mushroom	[123]
	6-methyl-5-hepten-2-one	[40]	fruity	[169]
OR52D1	acetophenone	[40]	vegetal	[157]
	anisole	[40]	phenolic	[197]
	benzaldehyde	[40]	almond	[131]
	benzothiazol	[40]	rubbery	[157]
	beta-ionone	[40]	violet	[120]
	butyl butyrate	[40]	pear	[198]
	butyric acid	[40]	cheesy	[123]
	caproic acid	[40]	goaty	[195]
	caprylic acid	[40]	fatty	[131]
	cinnamaldehyde	[40]	cinnamon	[116]
	citral	[40]	lemon	[125]
	citralva	[40]	citrus	[199]
	cyclohexanone	[40]	acetone	[200]
	decanal	[40]	waxy	[133]
	estragol	[40]	anise	[120]
	ethyl butyrate	[40]	strawberry	[131]
	ethyl caproate	[40]	ripe fruits	[131]

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
	ethyl heptanoate	[40]	overripe	[130]
	helional	[40]	floral	[146]
	heptanoic acid	[40]	cheesy	[194]
	isoamyl acetate	[40]	overripe banana	[171]
	isobutyric acid	[40]	cheese	[131]
	isovaleric acid	[40]	cheese	[123]
	lauric aldehyde	[40]	waxy	[133]
	methyl heptanoate	[40]	fruity	[147]
	methyl octanoate	[40]	green	[159]
	nonanal	[40]	orange	[123]
	nonanoic acid	[40]	coriander	[153]
	octanal	[40]	watermel on	[131]
	para-anisaldehyde	[40]	anise	[157]
	phenylmethanol	[40]	green	[174]
	propionic acid	[40]	vinegar	[194]
	safrole	[40]	spicy	[177]
	S-methylthio butanoate	[40]	cheesy	[178]
	thiazol	[40]	meaty	[179]
	trans-anethol	[40]	anise	[120]
	trans-cinnamic acid	[40]	pungent	[140]
OR56A1	decanal	[92]	waxy	[133]
OR56A4	decanal	[92]	waxy	[133]
OR56A5	undecanal	[42]	waxy	[148]
OR5A1	decanal	[92]	waxy	[133]
OR5A2	beta-ionone	[103]	violet	[120]
OR5AC2	beta-ionone	[103]	violet	[120]
	alpha-damascone	[166]	fruity	[139]
	eugenyl acetate	[166]	spicy	[119]
	fructone	[166]	fruity	[115]
	maltyl isobutyrate	[166]	fruity	[115]
	manzanate	[166]	apple	[115]
	vanilin	[166]	vanilla	[118]
OR5AN1	muscone	[10]	musky	[115]
OR5B17	eugenyl acetate	[166]	spicy	[119]

Table 3. Deorphanized ORs with corresponding agonists and their odors. Olfactory notes are represented by a colored square.

	■ animal ■ balsamic	■ woody ■ spicy	■ floral ■ sulfur	■ fruity ■ green
	floral ozone	[166]	ocean	[115]
	2-heptanone	[92]	banana	[125]
	benzene	[92]	ethereal	[200]
	beta-damascone	[92]	rose	[137]
	citral	[92]	lemon	[125]
OR5K1	eugenol	[92]	spicy	[116]
	eugenol methyl	[42, 92]	smoky	[120]
	limonene	[92]	lemon	[157]
	lyral	[92]	floral	[121]
	nonanal	[92]	orange	[149]
	propanal	[92]	pungent	[148]
	quinine	[92]	odorless	-
	(-)carvone	[39]	green	[125]
	(+)-carvone	[39, 92]	mint	[126]
OR5P3	1-heptanol	[39]	green	[131]
	1-hexanol	[39]	green	[131]
	4-hydroxycoumarin	[39, 92]	woodruff	[120]
	acetophenone	[39]	vegetal	[157]
	coumarin	[42, 92]	woodruff	[120]
	quinoline	[92]	coal tar	[148]
OR6P1	anisaldehyde	[42]	pungent	[201]
OR7C1	androstadienone	[42]	musky, urine	[99]
OR7D4	androstadienone	[42, 99]	musky, urine	[99]
	androstenone	[42, 99]	musky, urine	[99]
OR8B3	(+)-carvone	[42]	mint	[126]
OR8D1	sotolone	[42, 92]	caramel	-
OR8K3	(+)-menthol	[42, 92]	mint	[128]
	beta-damascenone	[92]	rose	[115]

A quick look at Table 3 clearly emphasizes the lack of a correlation between odors and a single olfactory receptor (see also Table 4). In fact, some receptors detect chemicals that are associated to the six general olfactory notes defined here. Specifically, broadly tuned OR1A1, OR1G1, OR2J2, OR2W1 and OR52D1 were shown to respond to very large sets of odorants belonging to different chemical families. In contrast, narrowly tuned receptors exhibit a very specific recognition for a given chemical, such as OR7D4 for androstenone and androstadienone. Although, similar molecules within a series are generally agonists for the same receptor ((+)) and (-)-carvone are equally detected by OR2W1), OR deorphanization allowed us to identify some receptors that are able to discriminate between very similar compounds. The differential perception between (S)-(-)-citronellol (rose) and (R)-(+)-citronellol (floral) could be partially due to a differential activation of OR1A2 with respect to OR1A1. However, OR1A2 has been shown to be weakly expressed in humans.[26, 41] Nevertheless, from a genetic point of view, the sequence homology between these two receptors is greater than 60%, emphasizing the discriminating power of our olfactory system even with receptors sharing a large sequence consensus.

The six main olfactory notes are not equally covered across the responses of the ORs (Table 4). Although floral and animal notes are represented in 24 out of 57 ORs studied so far, other notes are less present. The spicy note is spread over 39%, balsamic 35%, fruity 30%, green 28%, woody 18% and sulfur 16% of the studied ORs. Key food odorants are mostly associated with spicy, fruity, sulfur, green, and balsamic notes. Animal notes are related to body odors, and this category is also well investigated.

Table 4. Combinatorial code of OR activation for some odorants. An empty cell can either be associated with a lack of response or to a non-tested OR

hOR	animal	woody	floral	fruity	balsamic	spicy	sulfur	green
OR10A6	•							
OR10G3			•					
OR10G4					•	•		
OR10G7					•	•		
OR10G9					•	•	•	
OR10J5			•					
OR11A1	•	•						
OR11H4	•							
OR11H6	•							
OR11H7P	•							
OR1A1	•	•	•	•	•	•	•	•
OR1A2	•	•	•	•	•	•	•	•
OR1C1	•	•	•	•	•	•	•	•
OR1D2	•	•	•	•	•	•	•	•
OR1E3								
OR1G1	•	•	•	•	•	•	•	•
OR1L3	•	•	•	•	•	•	•	•
OR2A25	•							
OR2AG1				•				
OR2AT4	•	•						
OR2B11	•	•	•					
OR2B3	•	•	•					
OR2C1								
OR2G2								
OR2J2	•	•	•	•	•	•	•	•
OR2J3	•	•	•	•	•	•	•	•
OR2M2	•	•	•	•	•	•	•	•
OR2M4								
OR2M7								
OR2T10								
OR2T34								
OR2W1	•	•	•	•	•	•	•	•
OR3A1	•	•	•	•	•	•	•	•
OR4D6								
OR4D9								
OR4E2								
OR4Q3								
OR51E1	•							
OR51E2	•							
OR51L1	•							
OR52D1	•	•	•	•	•	•	•	•
OR56A1	•							
OR56A4	•							
OR56A5	•							
OR5A1	•							
OR5A2	•							
OR5AC2								
OR5AN1	•							
OR5B17								
OR5K1	•	•	•	•	•	•	•	•
OR5P3	•	•	•	•	•	•	•	•
OR6P1	•	•	•	•	•	•	•	•
OR7C1	•							
OR7D4	•							
OR8B3								
OR8D1								
OR8K3			•					

Table 5. Combinatorial code of OR activation for some odorants. A dot indicates that the OR was tested but didn't respond while an empty cell means ‘not tested’.

hOR	androstadienone	quinoline	cinnamaldehyde	eugenol	citral	nonanal	coumarin	ethyl vanillin	helional
OR10A6	•	•	•	•	•	•	•	•	•
OR10G3		•	•	•	•	•	•	•	•
OR10G4		•	•	•	•	•	•	•	•
OR10G7		•	•	•	•	•	•	•	•
OR10G9		•	•	•	•	•	•	•	•
OR1A1	•	•	•	•	•	•	•	•	•
OR1A2		•	•	•	•	•	•	•	•
OR1C1									
OR1G1									
OR2A25	•	•	•	•	•	•	•	•	•
OR2B11	•	•	•	•	•	•	•	•	•
OR2B3	•	•	•	•	•	•	•	•	•
OR2C1									
OR2G2									
OR2J2	•	•	•	•	•	•	•	•	•
OR2J3	•	•	•	•	•	•	•	•	•
OR2M2	•	•	•	•	•	•	•	•	•
OR2M4									
OR2M7									
OR2T10									
OR2T34									
OR2W1		•	•	•	•	•	•	•	•
OR3A1									•
OR4Q3		•	•	•	•	•	•	•	•
OR51L1	•	•	•	•	•	•	•	•	•
OR52D1		•	•	•	•	•	•	•	•
OR5K1		•	•	•	•	•	•	•	•
OR5P3		•	•	•	•	•	•	•	•
OR7C1	•	•	•	•	•	•	•	•	•
OR7D4	•	•	•	•	•	•	•	•	•

The tentative connection between odors and a human OR was recently made on the basis of a pharmacophore. OR1G1 would contribute to code waxy, fatty or rose odors within our brain.[164] Note that both OR1G1 and OR52D1 have also been associated with ‘fruity’, ‘sweet’ and ‘fat’ odors through an alternative bioinformatics approach,[202] but their chemical space is quite different (alcohols for 1G1 vs. acids for 52D1).[40] This is only partially in line with the results of Table 4, which highlights the broad tuning of these receptors (with OR1A1, OR2J2 and OR2W2) and hence their contribution to the detection of chemicals associated with all olfactory families. A detailed analysis focusing on the most studied chemicals reveals interesting points.

Table 5 includes examples of OR activation for nine odorants for which the data are well documented. Citral and nonanal share the ‘fruity’ note, and they exhibit a very similar combinatorial code with six ORs in common among the eight studied with also seven non-responsive common ORs, which is consistent with their common descriptors within the citrus family, *i.e.*, ‘lemon’ and ‘orange’, respectively.[123, 125] Cinnamaldehyde and eugenol belong to the spicy family, but they only share four common ORs, five non-responsive ORs and three differential responses for ORs studied with both molecules. Here, although their general family is spicy, they have very different specificities; the former is ‘cinnamon’, and the latter is ‘clove’.[116] Quinoline and androstadienone are within the same case (3 ORs in common, 3 non-responsive for both and 5 different activations) and in the same general family but have very different descriptor types (‘coal tar’, ‘close to indole’ vs. ‘musky’, ‘urinous’). Ethyl vanillin (‘balsamic’) and coumarin (‘woody’) show patterns that are far removed from those of the other, underlying that their related odors are associated to specific combinatorial codes. Notice that helional, although associated with the ‘floral’ family, shows a pattern similar to that of ‘fruity’ odorants (nonanal and citral), confirming that these descriptors are closely related. The combinatorial code that is presented remains rather simplistic, as it does not include the fact that some ORs recognize some odorants as inverse agonists. This code then expands to include negative responses, emphasizing its complexity. The role played by these inhibitory signals remains to be uncovered.[93]

Focusing on molecules specific to the fragrance community, it appears that they are poorly covered to date. Within the woody set, the most valuable families (sandalwood, vetiver, cedar or patchouli) are almost absent. A few sandalwood compounds have been investigated for their re-epithelialization through the activation of hOR2T4, which has been shown to be expressed within human keratinocytes.[183] Furthermore, other perfume-specific notes such as amber or marine are also lacking. This is actually counterintuitive with regard to the large investment in studies that are focused on the discovery of novel molecules representing these notes. The organic and analytical

chemist community working in the field of fragrances is indeed very active, notably through the use of organic synthesis and sometimes associated with olfactophore investigation.[37] One question is what is actually represented by these olfactophores, as they gather common chemical and structural features of odorants with the same odor. It is likely that, hidden within these olfactophores, information about the ORs involved in that odor's combinatorial code is present. Although the explicit role of ORs is not represented in those models, they remain our closest way to connect molecular structures to odors.

Odor is associated with a combinatorial code, and any model that attempts to predict this feature must rely on the differential activation of the ORs expressed within our nasal epithelium. Relationships between the odor and ORN spaces have been put forward based on a limited model comprised of 40 ORNs,[25] or by identifying common points in the activation of rat ORN by odorants within the sandalwood family.[203] As shown earlier, androstenone and OR7D4 might not represent the general rule. For many cases, a more elaborated pattern of OR activation is surely prevalent.

If the model is to mimic the response of an ORN, it must be focused on the predominant role of its expressed receptor. With regard to receptor chemoreception, the so-called “peri-receptors” events must be taken into consideration. Within the mucus, at least two types of proteins are expressed that undoubtedly contribute to the sense of smell. Although their role is still not totally understood, odorant binding proteins (OBPs) might act as scavengers that prevent ORs from becoming overstimulated in addition to their function of solubilizing odorants. A modulation of the OR response in presence of OBPs has already been put forward, but additional control experiments should be performed.[79, 80] Interestingly, among the proteins found in human mucus, glutathione-S-transferases (GSTs) are ideal candidates to modulate the odorant–receptor chemoreception. Within this family, two enzymes, GSTA1 and GSTP1, could play an important role.[204-206] The role of GSTs in human olfaction is supported by the fact that they have been demonstrated to be involved in insect olfaction and are present within the human mucus. Focusing on mammals, GST enzymes isolated from the rat olfactory epithelium show an activity related to odorant molecules. Those proteins could modify the chemical composition of the sniffed airflow, prior to the chemoreception of the OR in multiple ways. They could inactivate/metabolize/eliminate odorants to end/modify/prevent OR saturation in real time.[24] Multiple chemical stimulations could also occur even in the absence of an enzyme. The water-phase chemical equilibrium between an aldehyde and its 1,1-geminal diol derivative has been reported to be responsible for the activation of a receptor that was initially thought to respond specifically to aldehydes.[207] As discussed earlier, these effects are likely to be taken into account through *in vivo* deorphanization. This represents a clear step forward in connecting the chemical space with the perceptual space through neuronal activation.[11]

Regarding the fit between OR activation and odor description, it remains unknown whether tests on an exhaustive list of ORNs would provide a better fit or whether the system has reached its limit. This could also be because the odor space is more poorly defined than either the ORN or chemical spaces, which rely on more robust scientific descriptors. In principle, the odor description represents half of the SOR. It is likely that the description of the odor of a chemical by an individual is subjected to several personal, cultural[208] or emotional traits[209] that are independent of genetic characteristics. An efficient model will necessarily rely on an unambiguous description of the perceptual space, suggesting that linguistics will certainly play a major role. For example, can the descriptors ‘fresh’ and ‘cool’ be considered synonyms?[210] Another way to limit inter-individual differences would involve the use of a panel of individuals who describe chemicals with appropriate protocols towards a consensus, associating a prototypical odor with a single odorant.[211] Within this postgenomic era, the amount of information gained about the response of ORs, odor descriptors and human behaviors about odorant stimulation remain to be accurately connected. With the emergence of big-data, the use of elaborated large-scale data mining will certainly help in this project. It is even conceivable that with the continual drop in the price of genome sequencing, the emergence of personalized perfumes could appear earlier than thought. Certainly, large trans-disciplinary investigations remain to be developed, and computer scientists, chemists and biochemists, linguists, perfumers and flavorists and neuroscientists will continue gaining knowledge about structure-odor relationships on a rational physiological and genetic basis.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] G. Ohloff, E. Theimer **Fragrance Chemistry. The Science of the Sense of Smell**; 1982.
- [3] K.J. Rossiter, Structure-Odor Relationships, *Chem. Rev.* 96 (1996) 3201-3240.
- [4] A. Dravnieks **Atlas of odor character profiles**. Baltimore; 1985.
- [5] S. Arctander **Perfume and flavor materials of natural origin**. U.S.A.; 1960.
- [6] M. Chastrette, A. Elmouaffek, P. Sauvegrain, A multidimensional statistical study of similarities between 74 notes used in perfumery, *Chem. Sens.* 13 (1988) 295-305.
- [7] M. Pintore, C. Wechman, G. Sicard, M. Chastrette, N. Amaury, J.R. Chretien, Comparing the Information Content of Two Large Olfactory Databases, *J. Chem. Inf. Model.* 46 (2005) 32-38.
- [8] M. Boelens, H. Boelens, Some aspects of qualitative structure-odor relationships, *Perfumer & Flavorist* 28 (2003) 36-45.
- [9] P. Kraft, W. Eichenberger, Synthesis and Odor of Aliphatic Musk: Discovery of a New Class of Odorants, *Eur. J. Org. Chem.* 2004 (2004) 354-365.
- [10] M. Shirasu, K. Yoshikawa, Y. Takai, A. Nakashima, H. Takeuchi, H. Sakano, K. Touhara, Olfactory receptor and neural pathway responsible for highly selective sensing of musk odors, *Neuron* 81 (2014) 165-178.
- [11] T.S. McClintock, K. Adipietro, W.B. Titlow, P. Breheny, A. Walz, P. Mombaerts, H. Matsunami, In Vivo Identification of Eugenol-Responsive and Muscone-Responsive Mouse Odorant Receptors, *J. Neurosci.* 34 (2014) 15669-15678.
- [12] C. Delasalle, C.A. de March, U.J. Meierhenrich, H. Brevard, J. Golebiowski, N. Baldovini, Structure-Odor Relationships of Semisynthetic β -Santalol Analogs, *Chem. Biodiversity* 11 (2014) 1843-1860.
- [13] M. Boelens, H. Boelens, L. Van Gemert, Sensory properties of optical isomers, *Perfumer & Flavorist* 18 (1993) 1-16.
- [14] A. Chess, I. Simon, H. Cedar, R. Axel, Allelic inactivation regulates olfactory receptor gene expression, *Cell* 78 (1994) 823-834.
- [15] A. Magklara, S. Lomvardas, Stochastic gene expression in mammals: lessons from olfaction, *Trends Cell. Biol.* 23 (2013) 449-456.
- [16] A. Duchamp, M.F. Revial, A. Holley, P. Mac Leod, Odor discrimination by frog olfactory receptors, *Chem. Sens.* 1 (1974) 213-233.
- [17] E.H. Polak, Multiple profile-multiple receptor site model for vertebrate olfaction, *J. Theor. Biol.* 40 (1973) 469-484.
- [18] G. Glusman, I. Yanai, I. Rubin, D. Lancet, The Complete Human Olfactory Subgenome, *Genome Res.* 11 (2001) 685-702.
- [19] A. Matsui, Y. Go, Y. Niimura, Degeneration of Olfactory Receptor Gene Repertoires in Primates: No Direct Link to Full Trichromatic Vision, *Mol. Biol. Evol.* 27 (2010) 1192-1200.
- [20] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial Receptor Codes for Odors, *Cell* 96 (1999) 713-723.
- [21] S. Richardson: **What is Postgenomics?** In: *annual meeting of the 4S Annual Meeting: 11-25 2014; Crowne Plaza Cleveland City Center Hotel, Cleveland, OH*.
- [22] T.V. Getchell, F.L. Margolis, M.L. Getchell, Perireceptor and receptor events in vertebrate olfaction, *Prog. Neurobiol.* 23 (1984) 317-345.
- [23] N. Thiebaud, S. Veloso Da Silva, I. Jakob, G. Sicard, J. Chevalier, F. Ménétrier, O. Berdeaux, Y. Artur, J.-M. Heydel, A.-M. Le Bon, Odorant Metabolism Catalyzed by Olfactory Mucosal Enzymes Influences Peripheral Olfactory Responses in Rats, *PLoS ONE* 8 (2013) e59547.
- [24] J.-M. Heydel, A. Coelho, N. Thiebaud, A. Legendre, A.-M.L. Bon, P. Faure, F. Neiers, Y. Artur, J. Golebiowski, L. Briand, Odorant-Binding Proteins and Xenobiotic Metabolizing Enzymes: Implications in Olfactory Perireceptor Events, *Anat. Rec.* 296 (2013) 1333-1345.

- [25] Y. Furudono, Y. Sone, K. Takizawa, J. Hirono, T. Sato, Relationship between Peripheral Receptor Code and Perceived Odor Quality, *Chem. Sens.* 34 (2009) 151-158.
- [26] C. Verbeurgt, F. Wilkin, M. Tarabichi, F. Gregoire, J.E. Dumont, P. Chatelain, Profiling of Olfactory Receptor Gene Expression in Whole Human Olfactory Mucosa, *PLoS ONE* 9 (2014) e96333.
- [27] D. Pierron, N.G. Cortés, T. Letellier, L.I. Grossman, Current relaxation of selection on the human genome: Tolerance of deleterious mutations on olfactory receptors, *Mol. Phylogenetic Evol.* 66 (2013) 558-564.
- [28] G. Glusman, A. Bahar, D. Sharon, Y. Pilpel, J. White, D. Lancet, The olfactory receptor gene superfamily: data mining, classification, and nomenclature, *Mamm. Genome* 11 (2000) 1016-1023.
- [29] C. Flegel, S. Manteniotis, S. Osthold, H. Hatt, G. Gisselmann, Expression Profile of Ectopic Olfactory Receptors Determined by Deep Sequencing, *PLoS ONE* 8 (2013) e55368.
- [30] G. Drutel, J.M. Arrang, J. Diaz, C. Wisnewsky, K. Schwartz, J.C. Schwartz, Cloning of OL1, a putative olfactory receptor and its expression in the developing rat heart, *Recept. Channels* 3 (1995) 33-40.
- [31] M. Parmentier, F. Libert, S. Schurmans, S. Schiffmann, A. Lefort, D. Eggerickx, C. Ledent, C. Mollereau, C. Gerard, J. Perret *et al*, Expression of members of the putative olfactory receptor gene family in mammalian germ cells, *Nature* 355 (1992) 453-455.
- [32] P. Blache, L. Gros, G. Salazar, D. Bataille, Cloning and Tissue Distribution of a New Rat Olfactory Receptor-like (OL2), *Biochem. Biophys. Res. Commun.* 242 (1998) 669-672.
- [33] A. Malki, J. Fiedler, K. Fricke, I. Ballweg, M.W. Pfaffl, D. Krautwurst, Class I odorant receptors, TAS1R and TAS2R taste receptors, are markers for subpopulations of circulating leukocytes, *J. Leukoc. Biol.* 97 (2015) 533-545.
- [34] J.L. Pluznick, R.J. Protzko, H. Gevorgyan, Z. Peterlin, A. Sipos, J. Han, I. Brunet, L.-X. Wan, F. Rey, T. Wang *et al*, Olfactory receptor responding to gut microbiota-derived signals plays a role in renin secretion and blood pressure regulation, *Proc. Natl. Acad. Sci. USA* 110 (2013) 4410-4415.
- [35] N. Kang, J. Koo, Olfactory receptors in non-chemosensory tissues, *BMB Rep.* 45 (2012) 612-622.
- [36] S.R. Foster, E. Roura, W.G. Thomas, Extrasensory perception: Odorant and taste receptors beyond the nose and mouth, *Pharmacol. Ther.* 142 (2014) 41-61.
- [37] G. Ohloff, W. Pickenhagen, P. Kraft **Scent and Chemistry—The Molecular World of Odors**. Zürich: Weinheim; 2011.
- [38] R. Haddad, R. Khan, Y.K. Takahashi, K. Mori, D. Harel, N. Sobel, A metric for odorant comparison, *Nat. Methods* 5 (2008) 425-429.
- [39] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, J.D. Mainland, Odor Coding by Mammalian Receptor Repertoire, *Sci. Signal.* 2 (2009) ra9.
- [40] G. Sanz, C. Schlegel, J.-C. Pernollet, L. Briand, Comparison of Odorant Specificity of Two Human Olfactory Receptors from Different Phylogenetic Classes and Evidence for Antagonism, *Chem. Sens.* 30 (2005) 69-80.
- [41] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *J. Struct. Biol.* 159 (2007) 400-412.
- [42] J.D. Mainland, A. Keller, Y.R. Li, T. Zhou, C. Trimmer, L.L. Snyder, A.H. Moberly, K.A. Adipietro, W.L.L. Liu, H. Zhuang *et al*, The missense of smell: functional variability in the human odorant receptor repertoire, *Nat. Neurosci.* 17 (2014) 114-120.
- [43] D. Krautwurst, M. Kotthoff, in: C.J. Crasto (Eds.), *Olfactory Receptors, A Hit Map-Based Statistical Method to Predict Best Ligands for Orphan Olfactory Receptors: Natural Key Odorants Versus “Lock Picks”* 2013, pp. 85-97.
- [44] A. Dunkel, M. Steinhaus, M. Kotthoff, B. Nowak, D. Krautwurst, P. Schieberle, T. Hofmann, Nature’s Chemical Signatures in Human Olfaction: A Foodborne Perspective for Future Biotechnology, *Angew. Chem. Int. Ed. Engl.* 53 (2014) 7124-7143.

- [45] Y. Niimura, A. Matsui, K. Touhara, Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals, *Genome Res.* 24 (2014) 1485-1496.
- [46] H. Zhao, L. Ivic, J.M. Otaki, M. Hashimoto, K. Mikoshiba, S. Firestein, Functional Expression of a Mammalian Odorant Receptor, *Science* 279 (1998) 237-242.
- [47] J.R. Murrell, D.D. Hunter, An Olfactory Sensory Neuron Line, Odora, Properly Targets Olfactory Proteins and Responds to Odorants, *J. Neurosci.* 19 (1999) 8260-8270.
- [48] M.Q. Nguyen, Z. Zhou, C.A. Marks, N.J. Ryba, L. Belluscio, Prominent roles for odorant receptor coding sequences in allelic exclusion, *Cell* 131 (2007) 1009-1017.
- [49] P. Thomas, T.G. Smart, HEK293 cell line: A vehicle for the expression of recombinant proteins, *J. Pharmacol. Toxicol. Methods* 51 (2005) 187-200.
- [50] H. Zhuang, H. Matsunami, Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells, *Nat. Protoc.* 3 (2008) 1402-1413.
- [51] H. Saito, M. Kubota, R.W. Roberts, Q. Chi, H. Matsunami, RTP Family Members Induce Functional Expression of Mammalian Odorant Receptors, *Cell* 119 (2004) 679-691.
- [52] E. Shirokova, K. Schmiedeberg, P. Bedner, H. Niessen, K. Willecke, J.-D. Raguse, W. Meyerhof, D. Krautwurst, Identification of Specific Ligands for Orphan Olfactory Receptors G PROTEIN-DEPENDENT AGONISM AND ANTAGONISM OF ODORANTS, *J. Biol. Chem.* 280 (2005) 11807-11815.
- [53] A. Ukhakov, A. Shpenkova, M. Mergenov, R. Barto, [Primary-multiple carcinoma of the uterine cervix and stomach in an elderly woman], *Vestnik khirurgii imeni II Grekova* 165 (2005) 77-77.
- [54] B. Hsiao, K.B. Mihalak, S.E. Repicky, D. Everhart, A.H. Mederos, A. Malhotra, C.W. Luetje, Determinants of Zinc Potentiation on the α 4 Subunit of Neuronal Nicotinic Receptors, *Mol. Pharmacol.* 69 (2006) 27-36.
- [55] K.B. Mihalak, F.I. Carroll, C.W. Luetje, Varenicline is a partial agonist at α 4 β 2 and a full agonist at α 7 neuronal nicotinic receptors, *Mol. Pharmacol.* 70 (2006) 801-805.
- [56] D.J. Speca, D.M. Lin, P.W. Sorensen, E.Y. Isacoff, J. Ngai, A.H. Dittman, Functional Identification of a Goldfish Odorant Receptor, *Neuron* 23 (1999) 487-498.
- [57] C.H. Wetzel, M. Oles, C. Wellerdieck, M. Kuczkiwiak, G. Gisselmann, H. Hatt, Specificity and Sensitivity of a Human Olfactory Receptor Functionally Expressed in Human Embryonic Kidney 293 Cells andXenopus Laevis Oocytes, *J. Neurosci.* 19 (1999) 7426-7433.
- [58] C.H. Wetzel, H.-J. Behrendt, G. Gisselmann, K.F. Störkuhl, B. Hovemann, H. Hatt, Functional expression and characterization of a *Drosophila* odorant receptor in a heterologous cell system, *Proc. Natl. Acad. Sci. USA* 98 (2001) 9377-9380.
- [59] S. Katada, T. Nakagawa, H. Kataoka, K. Touhara, Odorant response assays for a heterologously expressed olfactory receptor, *Biochem. Biophys. Res. Commun.* 305 (2003) 964-969.
- [60] T. Abaffy, H. Matsunami, C.W. Luetje, Functional analysis of a mammalian odorant receptor subfamily, *J. Neurochem.* 97 (2006) 1506-1518.
- [61] I. Menashe, T. Abaffy, Y. Hasin, S. Goshen, V. Yahalom, C.W. Luetje, D. Lancet, Genetic Elucidation of Human Hyperosmia to Isovaleric Acid, *PLoS Biol.* 5 (2007) e284.
- [62] K. Raming, J. Krieger, J. Strotmann, I. Boekhoff, S. Kubick, C. Baumstark, H. Breer, Cloning and expression of odorant receptors, *Nature* 361 (1993) 353-356.
- [63] V. Matarazzo, O. Clot-Faybesse, B. Marcket, G. Guiraudie-Capraz, B. Atanasova, G. Devauchelle, M. Cerutti, P. Etiévant, C. Ronin, Functional Characterization of Two Human Olfactory Receptors Expressed in the Baculovirus Sf9 Insect Cell System, *Chem. Sens.* 30 (2005) 195-207.
- [64] L. Charlier, J. Topin, C. Ronin, S.K. Kim, W.A. Goddard, 3rd, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cell. Mol. Life Sci.* 69 (2012) 4205-4213.
- [65] J. Topin, C.A. de March, L. Charlier, C. Ronin, S. Antonczak, J. Golebiowski, Discrimination between Olfactory Receptor Agonists and Non-agonists, *Chem. Eur. J.* 20 (2014) 10227-10230.
- [66] Z. Peterlin, S. Firestein, M.E. Rogers, The state of the art of odorant receptor deorphanization: A report from the orphanage, *J. Gen. Physiol.* 143 (2014) 527-542.

- [67] K. King, H. Dohlman, J. Thorner, M. Caron, R. Lefkowitz, Control of yeast mating signal transduction by a mammalian beta 2-adrenergic receptor and Gs alpha subunit, *Science* 250 (1990) 121-123.
- [68] P. Sander, S. Grünwald, M. Bach, W. Haase, H. Reiländer, H. Michel, Heterologous Expression of the Human D2S Dopamine Receptor in Protease-Deficient *Saccharomyces cerevisiae* Strains, *Eur. J. Biochem.* 226 (1994) 697-705.
- [69] L.A. Price, E.M. Kajkowski, J.R. Hadcock, B.A. Ozenberger, M.H. Pausch, Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway, *Mol. Cell. Biol.* 15 (1995) 6188-6195.
- [70] N.E. David, M. Gee, B. Andersen, F. Naider, J. Thorner, R.C. Stevens, Expression and Purification of the *Saccharomyces cerevisiae* α-Factor Receptor (Ste2p), a 7-Transmembrane-segment G Protein-coupled Receptor, *J. Biol. Chem.* 272 (1997) 15553-15561.
- [71] J.R. Erickson, J.J. Wu, J.G. Goddard, G. Tigyi, K. Kawanishi, L.D. Tomei, M.C. Kiefer, Edg-2/Vzg-1 couples to the yeast pheromone response pathway selectively in response to lysophosphatidic acid, *J. Biol. Chem.* 273 (1998) 1506-1510.
- [72] I. Erlenbach, E. Kostenis, C. Schmidt, F.F. Hamdan, M.H. Pausch, J. Wess, Functional expression of M1, M3 and M5 muscarinic acetylcholine receptors in yeast, *J. Neurochem.* 77 (2001) 1327-1337.
- [73] J. Minic, M.A. Persuy, E. Godel, J. Aioun, I. Connerton, R. Salesse, E. Pajot-Augy, Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening, *FEBS J.* 272 (2005) 524-537.
- [74] V. Radhika, T. Proikas-Cezanne, M. Jayaraman, D. Onesime, J.H. Ha, D.N. Dhanasekaran, Chemical sensing of DNT by engineered olfactory yeast strain, *Nat. Chem. Biol.* 3 (2007) 325-330.
- [75] M. Spehr, G. Gisselmann, A. Poplawski, J.A. Riffell, C.H. Wetzel, R.K. Zimmer, H. Hatt, Identification of a Testicular Odorant Receptor Mediating Human Sperm Chemotaxis, *Science* 299 (2003) 2054-2058.
- [76] T. Strünker, N. Goodwin, C. Brenker, N.D. Kashikar, I. Weyand, R. Seifert, U.B. Kaupp, The CatSper channel mediates progesterone-induced Ca²⁺ influx in human sperm, *Nature* 471 (2011) 382-386.
- [77] C. Brenker, N. Goodwin, I. Weyand, N.D. Kashikar, M. Naruse, M. Krähling, A. Müller, U.B. Kaupp, T. Strünker, The CatSper channel: a polymodal chemosensor in human sperm, *Embo J.* 31 (2012) 1654-1665.
- [78] R. Neubig **Pharmacology of G Protein Coupled Receptors**, vol. 62: Academic Press; 2011.
- [79] J. Vidic, J. Grosclaude, R. Monnerie, M.A. Persuy, K. Badonnel, C. Baly, M. Caillol, L. Briand, R. Salesse, E. Pajot-Augy, On a chip demonstration of a functional role for Odorant Binding Protein in the preservation of olfactory receptor activity at high odorant concentration, *Lab on a chip* 8 (2008) 678-688.
- [80] H.J. Ko, S.H. Lee, E.H. Oh, T.H. Park, Specificity of odorant-binding proteins: a factor influencing the sensitivity of olfactory receptor-based biosensors, *Bioprocess Biosyst. Eng.* 33 (2010) 55-62.
- [81] H.J. Ko, T.H. Park, Enhancement of odorant detection sensitivity by the expression of odorant-binding protein, *Biosens. Bioelectron.* 23 (2008) 1017-1023.
- [82] V. Matarazzo, N. Zsürger, J.-C. Guillemot, O. Clot-Faybesse, J.-M. Botto, C.D. Farra, M. Crowe, J. Demaille, J.-P. Vincent, J. Mazella *et al*, Porcine Odorant-binding Protein Selectively Binds to a Human Olfactory Receptor, *Chem. Sens.* 27 (2002) 691-701.
- [83] K. Ukhanov, D. Brunert, E.A. Corey, B.W. Ache, Phosphoinositide 3-Kinase-Dependent Antagonism in Mammalian Olfactory Receptor Neurons, *J. Neurosci.* 31 (2011) 273-280.
- [84] X. Duan, E. Block, Z. Li, T. Connelly, J. Zhang, Z. Huang, X. Su, Y. Pan, L. Wu, Q. Chi, Crucial role of copper in detection of metal-coordinating odorants, *Proc. Natl. Acad. Sci. USA* 109 (2012) 3492-3497.

- [85] S. Sekharan, Mehmed Z. Ertem, H. Zhuang, E. Block, H. Matsunami, R. Zhang, Jennifer N. Wei, Y. Pan, Victor S. Batista, QM/MM Model of the Mouse Olfactory Receptor MOR244-3 Validated by Site-Directed Mutagenesis Experiments, *Biophys. J.* 107 (2014) L5-L8.
- [86] T. Pearce, J. Peacock, F. Aylward, D. Haisman, Catty odours in food: reactions between hydrogen sulphide and unsaturated ketones, *Chem. Ind.* 37 (1967) 1562.
- [87] H. Guth, Identification of Character Impact Odorants of Different White Wine Varieties, *J. Agric. Food Chem.* 45 (1997) 3022-3026.
- [88] R. Moncrieff **The Chemical Senses**. London; 1967.
- [89] B.D. Rubin, L.C. Katz, Optical Imaging of Odorant Representations in the Mammalian Olfactory Bulb, *Neuron* 23 (1999) 499-511.
- [90] R. Vincis, O. Gschwend, K. Bhaukaurally, J. Beroud, A. Carleton, Dense representation of natural odorants in the mouse olfactory bulb, *Nat. Neurosci.* 15 (2012) 537-539.
- [91] M. Meister, T. Bonhoeffer, Tuning and Topography in an Odor Map on the Rat Olfactory Bulb, *J. Neurosci.* 21 (2001) 1351-1360.
- [92] K.A. Adipietro, J.D. Mainland, H. Matsunami, Functional Evolution of Mammalian Odorant Receptors, *PLoS Genet.* 8 (2012) e1002821.
- [93] J. Reisert, Origin of basal activity in mammalian olfactory receptor neurons, *J. Gen. Physiol.* 136 (2010) 529-540.
- [94] Y. Oka, A. Nakamura, H. Watanabe, K. Touhara, An Odorant Derivative as an Antagonist for an Olfactory Receptor, *Chem. Sens.* 29 (2004) 815-822.
- [95] M.A. Chaput, F. El Mountassir, B. Atanasova, T. Thomas-Danguin, A.M. Le Bon, A. Perrut, B. Ferry, P. Duchamp-Viret, Interactions of odorants with olfactory receptors and receptor neurons match the perceptual dynamics observed for woody and fruity odorant mixtures, *Eur. J. Neurosci.* 35 (2012) 584-597.
- [96] T. Abaffy, A. Malhotra, C.W. Luetje, The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor: A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES, *J. Biol. Chem.* 282 (2007) 1216-1224.
- [97] J.F. McRae, J.D. Mainland, S.R. Jaeger, K.A. Adipietro, H. Matsunami, R.D. Newcomb, Genetic Variation in the Odorant Receptor OR2J3 Is Associated with the Ability to Detect the "Grassy" Smelling Odor, cis-3-hexen-1-ol, *Chem. Sens.* 37 (2012) 585-593.
- [98] E.A. Bremner, J.D. Mainland, R.M. Khan, N. Sobel, The Prevalence of Androstenone Anosmia, *Chem. Sens.* 28 (2003) 423-432.
- [99] A. Keller, H. Zhuang, Q. Chi, L.B. Vosshall, H. Matsunami, Genetic variation in a human odorant receptor alters odour perception, *Nature* 449 (2007) 468-472.
- [100] K. Lunde, B. Egelanddal, E. Skuterud, J.D. Mainland, T. Lea, M. Hersleth, H. Matsunami, Genetic Variation of an Odorant Receptor OR7D4 and Sensory Perception of Cooked Meat Containing Androstenone, *PLoS ONE* 7 (2012) e35259.
- [101] M. Blanch, N. Panella-Riera, P. Chevillon, M. Furnols, M. Gil, J. Gil, Z. Kallas, M. Oliver, Impact of consumer's sensitivity to androstenone on acceptability of meat from entire male pigs in three European countries: France, Spain and United Kingdom, *Meat Sci.* 90 (2012) 572-578.
- [102] A. Knaapila, L.-D. Hwang, A. Lysenko, F.F. Duke, B. Fesi, A. Khoshnevisan, R.S. James, C.J. Wysocki, M. Rhyu, M.G. Tordoff *et al*, Genetic Analysis of Chemosensory Traits in Human Twins, *Chem. Sens.* 37 (2012) 869-881.
- [103] S.R. Jaeger, J.F. McRae, C.M. Bava, M.K. Beresford, D. Hunter, Y. Jia, S.L. Chheang, D. Jin, M. Peng, J.C. Gamble *et al*, A Mendelian Trait for Olfactory Sensitivity Affects Odor Experience and Food Selection, *Curr. Biol.* 23 (2013) 1601-1605.
- [104] N.W. Eriksson, Shirley; Do, Chuong B.; Kiefer, Amy K.; Tung, Joyce Y.; Mountain, Joanna L.; Hinds, David A.; Francke, Uta, A genetic variant near olfactory receptor genes influences cilantro preference, *Flavour* 1 (2012) 22.
- [105] M.L. Pelchat, C. Bykowski, F.F. Duke, D.R. Reed, Excretion and Perception of a Characteristic Odor in Urine after Asparagus Ingestion: a Psychophysical and Genetic Study, *Chem. Sens.* 36 (2011) 9-17.

- [106] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Sci.* 13 (2004) 240-254.
- [107] L. Doszczak, P. Kraft, H.P. Weber, R. Bertermann, A. Triller, H. Hatt, R. Tacke, Prediction of perception: probing the hOR17-4 olfactory receptor model with silicon analogues of bourgeonal and lilial, *Angew. Chem. Int. Ed. Engl.* 46 (2007) 3367-3371.
- [108] J. Li, R. Haddad, S. Chen, V. Santos, C.W. Luetje, A broadly tuned mouse odorant receptor that detects nitrotoluenes, *J. Neurochem.* 121 (2012) 881-890.
- [109] L. Gelis, S. Wolf, H. Hatt, E. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angew. Chem. Int. Ed. Engl.* 51 (2012) 1274-1278.
- [110] C.A. de March, J. Golebiowski, A computational microscope focused on the sense of smell, *Biochimie* 107 (2014) 3-10.
- [111] P. Hummel, N. Vaidehi, W.B. Floriano, S.E. Hall, W.A. Goddard, Test of the Binding Threshold Hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93, *Protein Sci.* 14 (2005) 703-710.
- [112] S. Bavan, B. Sherman, C.W. Luetje, T. Abaffy, Discovery of Novel Ligands for Mouse Olfactory Receptor MOR42-3 Using an *<italic>In Silico</italic>* Screening Approach and *<italic>In Vitro</italic>* Validation, *PLoS ONE* 9 (2014) e92064.
- [113] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural Basis for a Broad But Selective Ligand Spectrum of a Mouse Olfactory Receptor: Mapping the Odorant-Binding Site, *J. Neurosci.* 25 (2005) 1806-1815.
- [114] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The Mouse Eugenol Odorant Receptor: Structural and Functional Plasticity of a Broadly Tuned Odorant Binding Pocket, *Biochemistry* 50 (2011) 843-853.
- [115] G.O. Brechbill **A reference book on fragrance ingredients**, vol. 30. New Jersey - USA: Frangrance Book Inc; 2006.
- [116] M.C. Diaz-Maroto, E. Guchu, L. Castro-Vázquez, C. De Torres, M.S. Perez-Coello, Aroma-active compounds of American, French, Hungarian and Russian oak woods, studied by GC-MS and GC-O, *Flavour Frag. J.* 23 (2008) 93-98.
- [117] S. Bader, M. Czerny, P. Eisner, A. Buettner, Characterisation of odour-active compounds in lupin flour, *J. Sci. Food Agric* 89 (2009) 2421-2427.
- [118] E. Sarrazin, D. Dubourdieu, P. Darriet, Characterization of key-aroma compounds of botrytized wines, influence of grape botrytization, *Food Chem.* 103 (2007) 536-545.
- [119] S.J. Ameh, N.N. Ibekwe, B.U. Ebeshi, Essential Oils in Ginger, Hops, Cloves, and Pepper Flavored Beverages—A Review, *J. Diet. Suppl.* (2014)
- [120] A. Zeller, M. Rychlik, Impact of estragole and other odorants on the flavour of anise and tarragon, *Flavour Frag. J.* 22 (2006) 105-113.
- [121] C. Anselmi, M. Centini, A. Sega, E. Napolitano, P. Pelosi, C. Scesa, Synthesis and odour properties of floral smelling compounds, *Int. J. Cosmetic Sci.* 18 (1996) 67-74.
- [122] E. Polak, D. Trotier, E. Baliguet, Odor similarities in structurally related odorants, *Chem. Sens.* 3 (1978) 369-380.
- [123] J.-L. Berdagué, P. Tournayre, S. Cambou, Novel multi-gas chromatography-olfactometry device and software for the identification of odour-active compounds, *J. Chromatogr. A* 1146 (2007) 85-92.
- [124] L. Jirovetz, G. Buchbauer, A. Stoyanova, A. Balinova, Z. Guangjiun, M. Xihan, Solid phase microextraction/gas chromatographic and olfactory analysis of the scent and fixative properties of the essential oil of Rosa damascena L. from China, *Flavour Frag. J.* 20 (2005) 7-12.
- [125] Y. Wang, C. Finn, M.C. Qian, Impact of growing environment on Chickasaw blackberry (*Rubus L.*) aroma evaluated by gas chromatography olfactometry dilution analysis, *J. Agric. Food Chem.* 53 (2005) 3563-3571.
- [126] H.-S. Choi, Aroma Evaluation of an Aquatic Herb, Changpo (*Acorus calamus* Var. *angustatus* Bess), by AEDA and SPME, *J. Agric. Food Chem.* 52 (2004) 8099-8104.

- [127] L. Jirovetz, G. Buchbauer, A.S. Stoyanova, E.V. Georgiev, S.T. Damianova, Composition, Quality Control, and Antimicrobial Activity of the Essential Oil of Long-Time Stored Dill (*Anethum graveolens* L.) Seeds from Bulgaria, *J. Agric. Food Chem.* 51 (2003) 3854-3857.
- [128] H.-S. Choi, Character Impact Odorants of Citrus Hallabong [(*C. unshiu* Marcov * *C. sinensis* Osbeck) * *C. reticulata* Blanco] Cold-Pressed Peel Oil, *J. Agric. Food Chem.* 51 (2003) 2687-2692.
- [129] T. Yamamoto, A. Shimada, T. Ohmoto, H. Matsuda, M. Ogura, T. Kanisawa, Olfactory study on optically active citronellyl derivatives, *Flavour Frag. J.* 19 (2004) 121-133.
- [130] A. Elston, J. Lin, R. Rouseff, Determination of the role of valencene in orange oil as a direct contributor to aroma quality, *Flavour Frag. J.* 20 (2005) 381-386.
- [131] L. Zea, L. Moyano, J.A. Moreno, M. Medina, Aroma series as fingerprints for biological ageing in fino sherry-type wines, *J. Sci. Food Agric* 87 (2007) 2319-2326.
- [132] S. Ishizaki, T. Tachihara, H. Tamura, T. Yanai, T. Kitahara, Evaluation of odour-active compounds in roasted shrimp (*Sergia lucens* Hansen) by aroma extraction dilution analysis, *Flavour Frag. J.* 20 (2005) 562-566.
- [133] S. Sellie, C. Rannou, C. Prost, J. Robin, T. Serot, Characterization of Aroma-active Compounds in Rainbow Trouts (*Oncorhynchus mykiss*) Eliciting an Off-Odor, *J. Agric. Food Chem.* 54 (2006) 9496-9502.
- [134] M.S. Madruga, J.S. Elmore, M.J. Oruna-Concha, D. Balagiannis, D.S. Mottram, Determination of some water-soluble aroma precursors in goat meat and their enrolment on flavour profile of goat meat, *Food Chem.* 123 (2010) 513-520.
- [135] K. Herbrand, F.J. Hammerschmidt, S. Brennecke, M. Liebig, G. Lösing, C.O. Schmidt, I. Gatfield, G. Krammer, H.-J. Bertram, Identification of Allyl Esters in Garlic Cheese, *J. Agric. Food Chem.* 55 (2007) 7874-7878.
- [136] M.W. Cheong, S.Q. Liu, W. Zhou, P. Curran, B. Yu, Chemical composition and sensory profile of pomelo (*Citrus grandis* (L.) Osbeck) juice, *Food Chem.* 135 (2012) 2505-2513.
- [137] D. Gocmen, O. Gurbuz, R. Rouseff, J. Smoot, A. Fatih Dagdelen, Gas chromatographic-olfactometric characterization of aroma active compounds in sun-dried and vacuum-dried tarhana, *Eur. Food Res. Technol.* 218 (2004) 573-578.
- [138] T. Braun, P. Voland, L. Kunz, C. Prinz, M. Gratzl, Enterochromaffin Cells of the Human Gut: Sensors for Spices and Odorants, *Gastroenterology* 132 (2007) 1890-1901.
- [139] G. Mosciano, M. Fasano, J. Michalski, S. Sadural, Organoleptic characteristics of flavor materials, *Perfumer & Flavorist* 16 (1991) 45-47.
- [140] C. Osorio, M. Alarcon, C. Moreno, A. Bonilla, J. Barrios, C. Garzon, C. Duque, Characterization of Odor-Active Volatiles in Champa (*Campomanesia lineatifolia* R. & P.), *J. Agric. Food Chem.* 54 (2006) 509-516.
- [141] S. Harikedua, C.H. Wijaya, D. Adawiyah, Relationship between sensory attributes of bakasang (a traditional Indonesian fermented fish product) and its physicochemical properties, *Fish. Sci.* 78 (2012) 187-195.
- [142] U. Ravid, R. Ikan, New syntheses in dihydrojasnone series, *J. Org. Chem.* 39 (1974) 2637-2639.
- [143] G. Mosciano, The Creative Flavorist: Advanced Biotech, *Perfumer & Flavorist* 23 (1998) 49.
- [144] B. d'Acampora Zellner, A. Casilli, P. Dugo, G. Dugo, L. Mondello, Odour fingerprint acquisition by means of comprehensive two-dimensional gas chromatography-olfactometry and comprehensive two-dimensional gas chromatography/mass spectrometry, *J. Chromatogr. A* 1141 (2007) 279-286.
- [145] M. Sawamura, Y. Onishi, Y. Ikemoto, N.T.M. Tu, N.T.L. Phi, Characteristic odour components of bergamot (*Citrus bergamia* Risso) essential oil, *Flavour Frag. J.* 21 (2006) 609-615.
- [146] D. Enders, M. Backes, First asymmetric synthesis of both enantiomers of Tropional® and their olfactory evaluation, *Tetrahedron: Asymmetry* 15 (2004) 1813-1817.
- [147] G. Mosciano, M. Fasano, B. Allen, J. Michalski, S. Sadural, Organoleptic characteristics of flavor materials, *Perfumer & Flavorist* 15 (1990) 35-38.
- [148] G. Mosciano, Organoleptic characteristics of flavor materials, *Perfumer & Flavorist* 26 (2001) 82-85.

- [149] S. Huanlu, K.R. Cadwallader, T.K. Singh, Odour-active compounds of Jinhua ham, Flavour Frag. J. 23 (2008) 1-6.
- [150] Q. Guo, X. Li, J. Yu, H. Zhang, Y. Zhang, M. Yang, N. Lu, D. Zhang, Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry for screening of potent swampy/septic odor causing compounds in two drinking water sources of China, Anal. Methods (2015)
- [151] S.M. Fors, B.K. Olofsson, Alkylpyrazines, volatiles formed in the Maillard reaction. II. Sensory properties of five alkylpyrazines, Chem. Sens. 11 (1986) 65-77.
- [152] D. Maßberg, A. Simon, D. Häussinger, V. Keitel, G. Gisselmann, H. Conrad, H. Hatt, Monoterpene (-)-citronellal affects hepatocarcinoma cell signaling via an olfactory receptor, Arch. Biochem. Biophys. 566 (2015) 100-109.
- [153] R. Iraqi, C. Vermeulen, A. Benzekri, A. Bouseta, S. Collin, Screening for Key Odorants in Moroccan Green Olives by Gas Chromatography-Olfactometry/Aroma Extract Dilution Analysis, J. Agric. Food Chem. 53 (2005) 1179-1184.
- [154] A. Triller, E.A. Boulden, A. Churchill, H. Hatt, J. Englund, M. Spehr, C.S. Sell, Odorant-Receptor Interactions and Odor Percept: A Chemical Perspective, Chem. Biodiversity 5 (2008) 862-886.
- [155] G. Mosciano, M. Fasano, J. Cassidy, K. Connelly, P. Mazeiko, A. Montenegro, B. ALLEN, J. MICHALSKI, S. SADURAL, Organoleptic characteristics of flavor materials, Perfumer & Flavorist 19 (1994) 51.
- [156] B.L. Cook, D. Steuerwald, L. Kaiser, J. Graveland-Bikker, M. Vanberghem, A.P. Berke, K. Herlihy, H. Pick, H. Vogel, S. Zhang, Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4, Proc. Natl. Acad. Sci. USA 106 (2009) 11925-11930.
- [157] V. Varlet, C. Knockaert, C. Prost, T. Serot, Comparison of Odor-Active Volatile Compounds of Fresh and Smoked Salmon, J. Agric. Food Chem. 54 (2006) 3391-3401.
- [158] G. Eyres, J.-P. Dufour, G. Hallifax, S. Sotheeswaran, P.J. Marriott, Identification of character-impact odorants in coriander and wild coriander leaves using gas chromatography-olfactometry (GCO) and comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry (GC * GC-TOFMS), J. Sep. Sci. 28 (2005) 1061-1074.
- [159] C. Varming, M.A. Petersen, L. Poll, Comparison of isolation methods for the determination of important aroma compounds in black currant (*Ribes nigrum* L.) juice, using nasal impact frequency profiling, J. Agric. Food Chem. 52 (2004) 1647-1652.
- [160] C. Prost, A. Hallier, M. Cardinal, T. Serot, P. Courcoux, Effect of storage time on raw sardine (*Sardina pilchardus*) flavor and aroma quality, J. Food Sci. 69 (2004) S198-S204.
- [161] M.T. Morales, G. Luna, R. Aparicio, Comparative study of virgin olive oil sensory defects, Food Chem. 91 (2005) 293-301.
- [162] H. Abe, S. Kanaya, Y. Takahashi, S.-I. Sasaki, Extended studies of the automated odor-sensing system based on plural semiconductor gas sensors with computerized pattern recognition techniques, Anal. Chim. Acta 215 (1988) 155-168.
- [163] M. Chastrette, D. Cretin, A. El, Structure–Odor Relationships: Using Neural Networks in the Estimation of Camphoraceous or Fruity Odors and Olfactory Thresholds of Aliphatic Alcohols, J. Chem. Inform. Comput. Sci. 36 (1996) 108-113.
- [164] G. Sanz, T. Thomas-Danguin, E.H. Hamdani, C. Le Poupon, L. Briand, J.-C. Pernolle, E. Guichard, A. Tromelin, Relationships Between Molecular Structure and Perceived Odor Quality of Ligands for a Human Olfactory Receptor, Chem. Sens. 33 (2008) 639-653.
- [165] B. Lorrain, J. Ballester, T. Thomas-Daguin, J. Blanquet, J.M. Meunier, Y. Le Fur, Selection of potential impact odorants and sensory validation of their importance in typical Chardonnay wines, J. Agric. Food Chem. 54 (2006) 3973-3981.
- [166] D. Gonzalez-Kristeller, J.B.P. do Nascimento, P.A.F. Galante, B. Malnic, Identification of agonists for a group of human odorant receptors, Front. Pharmacol. 6 (2015)
- [167] W. Fan, M.C. Qian, Headspace Solid Phase Microextraction and Gas Chromatography-Olfactometry Dilution Analysis of Young and Aged Chinese \"Yanghe Daqu\" Liquors, J. Agric. Food Chem. 53 (2005) 7931-7938.

- [168] T. Kishimoto, A. Wanikawa, K. Kono, K. Shibata, Comparison of the Odor-Active Compounds in Unhopped Beer and Beers Hopped with Different Hop Varieties, *J. Agric. Food Chem.* 54 (2006) 8855-8861.
- [169] G. Mosciano, Organoleptic characteristics of flavor materials, *Perfumer & Flavorist* 22 (1997) 69-72.
- [170] C.H. Wijaya, D. Ulrich, R. Lestari, K. Schippel, G. Ebert, Identification of Potent Odorants in Different Cultivars of Snake Fruit [Salacca zalacca (Gaert.) Voss] Using Gas Chromatography-Olfactometry, *J. Agric. Food Chem.* 53 (2005) 1637-1641.
- [171] M.J. Gomez-Miguez, J.F. Cacho, V. Ferreira, I.M. Vicario, F.J. Heredia, Volatile components of Zalema white wines, *Food Chem.* 100 (2007) 1464-1473.
- [172] O. Gonzalez-Rios, M.L. Suarez-Quiroz, R. Boulanger, M. Barel, B. Guyot, J.-P. Guiraud, S. Schorr-Galindo, Impact of "ecological" post-harvest processing on the volatile fraction of coffee beans: I. Green coffee, *J. Food Comp. Anal.* 20 (2007) 289-296.
- [173] C.K. Iversen, H.B. Jakobsen, C.-E. Olsen, Aroma Changes during Black Currant (*Ribes nigrum* L.) Nectar Processing, *J. Agric. Food Chem.* 46 (1998) 1132-1136.
- [174] R.F.A. Moreira, C.A.B. de Maria, Investigation of the aroma compounds from headspace and aqueous solution from the cambara (*Gochnatia velutina*) honey, *Flavour Frag. J.* 20 (2005) 13-17.
- [175] D. McGinty, C.S. Letizia, A.M. Api, Fragrance material review on piperonyl acetate, *Food Chem. Toxicol.* 50, Supplement 2 (2012) S358-S362.
- [176] H. Masuda, S. Mihara, Olfactive properties of alkylpyrazines and 3-substituted 2-alkylpyrazines, *J. Agric. Food Chem.* 36 (1988) 584-587.
- [177] Y. Lorjaroenphon, K.R. Cadwallader, Characterization of Typical Potent Odorants in Cola-Flavored Carbonated Beverages by Aroma Extract Dilution Analysis, *J. Agric. Food Chem.* 63 (2014) 769-775.
- [178] X. Du, M. Song, R. Rouseff, Identification of New Strawberry Sulfur Volatiles and Changes during Maturation, *J. Agric. Food Chem.* 59 (2011) 1293-1300.
- [179] Q. Xu, J. Liu, H. Song, T. Zou, Y. Liu, S. Zhang, Formation mechanism of volatile and non-volatile compounds in peptide-xylose Maillard reaction, *Food Res. Int.* 54 (2013) 683-690.
- [180] M. Miyazawa, J. Kawata, Identification of the main aroma compounds in dried seeds of *Brassica hirta*, *J. Nat. Med.* 60 (2006) 89-92.
- [181] E.M. Neuhaus, A. Mashukova, W. Zhang, J. Barbour, H. Hatt, A Specific Heat Shock Protein Enhances the Expression of Mammalian Olfactory Receptor Proteins, *Chem. Sens.* 31 (2006) 445-452.
- [182] A. Mashukova, M. Spehr, H. Hatt, E.M. Neuhaus, β -Arrestin2-Mediated Internalization of Mammalian Odorant Receptors, *J. Neurosci.* 26 (2006) 9902-9912.
- [183] D. Busse, P. Kudella, N.-M. Gruning, G. Gisselmann, S. Stander, T. Luger, F. Jacobsen, L. Steinbräser, R. Paus, P. Gkogkolou et al, A Synthetic Sandalwood Odorant Induces Wound-Healing Processes in Human Keratinocytes via the Olfactory Receptor OR2AT4, *J. Invest. Dermatol.* 134 (2014) 2823-2832.
- [184] G. Mosciano, The Creative Flavorist: Aldrich Chemical, *Perfumer & Flavorist* 24 (1999) 10-13.
- [185] M.-F. Hérent, S. Collin, P. Pelosi, Affinities of Nutty and Green-smelling Pyrazines and Thiazoles to Odorant-binding Proteins, in Relation with their Lipophilicity, *Chem. Sens.* 20 (1995) 601-608.
- [186] V. Kraujalytė, E. Leitner, P.R. Venskutonis, Chemical and sensory characterisation of aroma of *Viburnum opulus* fruits by solid phase microextraction-gas chromatography-olfactometry, *Food Chem.* 132 (2012) 717-723.
- [187] S. Antoniotti, N. Alezra, X. Fernandez, E. Duñach, Catalytic epoxide oxidation: a novel access to flavouring and odorant α -diketones, *Flavour Frag. J.* 19 (2004) 373-381.
- [188] D.S. Garruti, M.R.B. Franco, M.A.A.P. da Silva, N.S. Janzanti, G.L. Alves, Assessment of aroma impact compounds in a cashew apple-based alcoholic beverage by GC-MS and GC-olfactometry, *LWT - Food Science and Technology* 39 (2006) 373-378.
- [189] G. Mosciano, M. Fasano, J. Cassidy, K. Connelly, P. Mazeiko, A. Montenegro, J. Michalski, S. Sadural, Organoleptic characteristics of flavor materials, *Perfumer & Flavorist* 20 (1995) 49-51.

- [190] E.K. Mo, C.K. Sung, Phenylethyl alcohol (PEA) application slows fungal growth and maintains aroma in strawberry, *Postharvest Biol. Technol.* 45 (2007) 234-239.
- [191] A. Plotto, M.R. McDaniel, J.P. Mattheis, Characterization of Changes in 'Gala' Apple Aroma during Storage Using Osme Analysis, a Gas Chromatography-Olfactometry Technique, *J. Am. Soc. Hortic. Sci.* 125 (2000) 714-722.
- [192] V. Jacquier, H. Pick, H. Vogel, Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds, *J. Neurochem.* 97 (2006) 537-544.
- [193] Y. Fujita, T. Takahashi, A. Suzuki, K. Kawashima, F. Nara, R. Koishi, Deorphanization of Dresden G Protein-Coupled Receptor for an Odorant Receptor, *J. Recept. Sig. Transd.* 27 (2007) 323-334.
- [194] L. Cai, J.A. Koziel, Y.-C. Lo, S.J. Hoff, Characterization of volatile organic compounds and odorants associated with swine barn particulate matter using solid-phase microextraction and gas chromatography-mass spectrometry-olfactometry, *J. Chromatogr. A* 1102 (2006) 60-72.
- [195] D. Komes, D. Ulrich, T. Lovric, Characterization of odor-active compounds in Croatian Rhine Riesling wine, subregion Zagorje, *Eur. Food Res. Technol.* 222 (2006) 1-7.
- [196] C. Derail, T. Hofmann, P. Schieberle, Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure, *J. Agric. Food Chem.* 47 (1999) 4742-4745.
- [197] A.J. MacLeod, N. Gonzales de Troconis, Volatile flavor components of sapodilla fruit (*Achras sapota* L.), *J. Agric. Food Chem.* 30 (1982) 515-517.
- [198] E. Mehinagic, G. Royer, R. Symoneaux, F. Jourjon, C. Prost, Characterization of Odor-Active Volatiles in Apples: Influence of Cultivars and Maturity Stage, *J. Agric. Food Chem.* 54 (2006) 2678-2687.
- [199] A.P.S. Narula, The Search for New Fragrance Ingredients for Functional Perfumery, *Chem. Biodiversity* 1 (2004) 1992-2000.
- [200] J.E. Amoore, D. Venstrom, Sensory Analysis of Odor Qualities in Terms of the Stereochemical Theory, *J. Food Sci.* 31 (1966) 118-128.
- [201] B. Siegmund, B. Pöllinger-Zierler, Odor Thresholds of Microbially Induced Off-Flavor Compounds in Apple Juice, *J. Agric. Food Chem.* 54 (2006) 5984-5989.
- [202] K. Audouze, A. Tromelin, A.M. Le Bon, C. Belloir, R.K. Petersen, K. Kristiansen, S. Brunak, O. Tabouret, Identification of Odorant-Receptor Interactions by Global Mapping of the Human Odorome, *PLoS ONE* 9 (2014) e93037.
- [203] S. Bieri, K. Monastyrskaya, B. Schilling, Olfactory Receptor Neuron Profiling using Sandalwood Odorants, *Chem. Sens.* 29 (2004) 483-487.
- [204] B. Casado, L.K. Pannell, P. Iadarola, J.N. Baraniuk, Identification of human nasal mucous proteins using proteomics, *Proteomics* 5 (2005) 2949-2959.
- [205] H. Débat, C. Eloit, F. Blon, B. Sarazin, C. Henry, J.-C. Huet, D. Trotier, J.-C. Pernollet, Identification of Human Olfactory Cleft Mucus Proteins Using Proteomic Analysis, *J. Proteome Res.* 6 (2007) 1985-1996.
- [206] P.V. Tomazic, R. Birner-Gruenberger, A. Leitner, B. Darnhofer, S. Spoerk, D. Lang-Loidolt, Apolipoproteins have a potential role in nasal mucus of allergic rhinitis patients: A proteomic study, *Laryngoscope* (2014)
- [207] Y. Li, Z. Peterlin, J. Ho, T. Yarnitzky, M.T. Liu, M. Fichman, M.Y. Niv, H. Matsunami, S. Firestein, K. Ryan, Aldehyde Recognition and Discrimination by Mammalian Odorant Receptors via Functional Group-Specific Hydration Chemistry, *ACS Chem. Biol.* 9 (2014) 2563-2571.
- [208] S. Ayabe-Kanamura, I. Schicker, M. Laska, R. Hudson, H. Distel, T. Kobayakawa, S. Saito, Differences in Perception of Everyday Odors: a Japanese-German Cross-cultural Study, *Chem. Sens.* 23 (1998) 31-38.
- [209] C. Ferenzi, A. Schirmer, S.C. Roberts, S. Delplanque, C. Porcherot, I. Cayeux, M.-I. Velasco, D. Sander, K.R. Scherer, D. Grandjean, Affective dimensions of odor perception: A comparison between Swiss, British, and Singaporean populations, *Emotion* 11 (2011) 1168-1181.

PARTIE 1 : Les mécanismes de la perception des odeurs

Article 1 – de March et al. FFJ 30 (2015) 342-361

[210] M. Zarzo, What is a Fresh Scent in Perfumery? Perceptual Freshness is Correlated with Substantivity, Sensors (Basel) 13 (2013) 463-483.

[211] M. Zarzo, D. Stanton, Understanding the underlying dimensions in perfumers' odor perception space as a basis for developing meaningful odor maps, Atten. Percept. Psychophys. 71 (2009) 225-247.

Article 2 - Les étapes moléculaires de la perception des odeurs décrites par les approches de modélisation moléculaire

La compréhension des comportements des protéines impliquées dans la perception des odeurs semble incontournable dans l'établissement des relations structure-odeur. Au-delà des ROs, un autre type de protéine est également suspecté de jouer un rôle important dans l'olfaction. Cette protéine, l'OBP (Odorant Binding Protein), a été découverte dans le mucus olfactif et a une grande affinité pour les molécules odorantes. Comprendre les mécanismes moléculaires de l'olfaction nécessite au minimum l'étude de ces deux protagonistes vis-à-vis des odorants. L'évolution des outils informatiques permet aujourd'hui d'aborder cette étude sous un angle computationnel. Plus en détail, la modélisation moléculaire permet de décrire la structure et le comportement dynamique des protéines avec une précision au niveau atomique, formant un « microscope computationnel ». Ces outils, basés sur la chimie théorique, permettent d'obtenir des modèles de protéines représentant chacun de ses atomes explicitement (Figure 1) et de les voir évoluer au cours du temps selon les lois de la physique. On parle de mécanique et de dynamique moléculaires. Ce « microscope » permet d'observer le comportement dynamique d'objets moléculaires afin de tenter d'identifier le mode de fonctionnement des ROs et des OBPs.

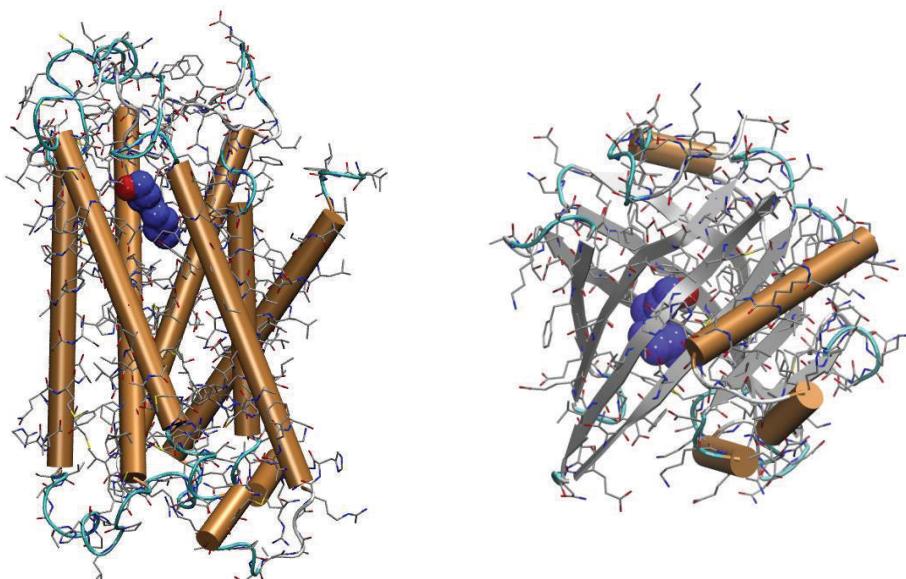


Figure 1. Gauche) modèle tridimensionnel du ROi7 de la souris. Elle a été obtenue par reconstruction par homologie et est liée à une molécule d'octanol. Droite) Structure expérimentale de l'OBP 3 du rat liée au décanol. Pour les deux modèles, la structure secondaire est symbolisée en cartoon (hélices en beige, feuilllets en gris et coudes en cyan). Les chaînes latérales des acides aminés sont représentées par des lignes (les atomes de carbone sont en gris, les atomes d'azote en bleu et les atomes d'oxygène en rouge). Le ligand est dans les deux cas présenté sous la forme de son volume de van der Waals en bleu foncé.

Les OBP dans les mécanismes de l'olfaction

Les OBP appartiennent à la famille des lipocalines. Elles sont exprimées dans le mucus olfactif qui recouvre notre épithélium olfactif. Ces protéines possèdent une forme rappelant celle d'un tonneau et sont composées d'une cavité de liaison hydrophobe et d'une surface extérieure hydrophile. Elles aideraient à la solubilisation des molécules odorantes, souvent hydrophobes, dans le mucus olfactif composé principalement d'eau. Solubilisent-elles les odorants pour les transporter jusqu'au ROs, ou pour nettoyer le mucus olfactif ? Entrent-elles en interaction avec les ROs ? Leur implication exacte n'est pour l'instant pas clairement établie.

Les ROs, la pierre angulaire de la perception des odeurs.

Les ROs interagissent avec les molécules odorantes et transforment le signal chimique de ces composés en signal électrique transmis à notre cerveau (que nous appelons une odeur). Nos gènes expriment environ 400 types de ROs fonctionnels. Ils sont majoritairement localisés dans nos neurones olfactifs mais ils peuvent également être exprimés de façon dite « ectopique » dans d'autres organes. Un odorant est capable d'activer plusieurs ROs et un RO peut être activé par plusieurs odorants. La perception de l'odeur d'une molécule résulte du code combinatoire d'activation de ROs qui lui est associé. Aussi, chacun de ces ROs peut avoir des comportements différents par rapport à l'espace d'odorants. Certains sont dits à large spectre car ils répondent à de nombreuses molécules, d'autres sont à spectre restreint car ils s'activent de façon spécifique à seulement quelques odorants. La compréhension des mécanismes d'activation, de liaison, de sélectivité est donc cruciale pour tenter de prédire le code combinatoire associé à une molécule sur la base de sa structure. Toutefois, aucune structure expérimentale de RO n'est connue à ce jour et la modélisation moléculaire s'impose comme un outil privilégié pour répondre à ces questions.

Comment prédire la structure de ces protéines ?

La reconstruction d'une protéine dont la structure n'a pas été élucidée expérimentalement peut se faire principalement selon deux protocoles. Les méthodes dites *ab initio* prennent pour point de départ la séquence d'acides aminés de la protéine cible (sa structure primaire). Elles prédisent la structure secondaire de celle-ci grâce aux propriétés de ses acides aminés. En appliquant les lois de la physique, il est possible de déduire la structure tertiaire la plus probable de la protéine.

La reconstruction dite « par homologie » consiste quant à elle à s'inspirer des structures expérimentalement connues de protéines proches. Dans le cas des OBP, la structure de quelques-unes d'entre elles est disponible, facilitant la reconstruction par homologie. Concernant les ROs, aucune structure expérimentale n'est connue à ce jour. La reconstruction par homologie se fait donc

par la prise en compte de structures de protéines appartenant à la même famille, les Récepteurs Couplés aux Protéines G.

Pour ces deux protocoles, l'alignement entre la séquence de la protéine cible et celles d'autres protéines de la même famille reste le point crucial. Un décalage même infime de cet alignement aura des conséquences sur le modèle final. On comprend donc facilement l'intérêt de concevoir un alignement de séquences pertinent et si possible validé par des méthodes expérimentales (de mutagénèse dirigée par exemple). Cet aspect a été développé dans ces travaux de recherche et sera présenté ultérieurement.

Quelle sont les performances des méthodes de chimie théorique ?

La modélisation moléculaire permet d'observer et d'analyser les systèmes moléculaires. Les méthodes les plus couramment utilisées dans l'étude des ROs et des OBPs sont le docking et la dynamique moléculaire. Le docking consiste, par exemple, à identifier la position d'une molécule odorante dans la cavité de liaison d'une protéine. La position la plus stable de cette molécule en fonction des propriétés de la cavité est prédite sur la base de leur énergie d'interaction. La dynamique moléculaire permet quant à elle de faire évoluer un système moléculaire grâce aux lois de la mécanique du point, ou mécanique Newtonienne. Dans ce cas le comportement dynamique du système peut être observé.

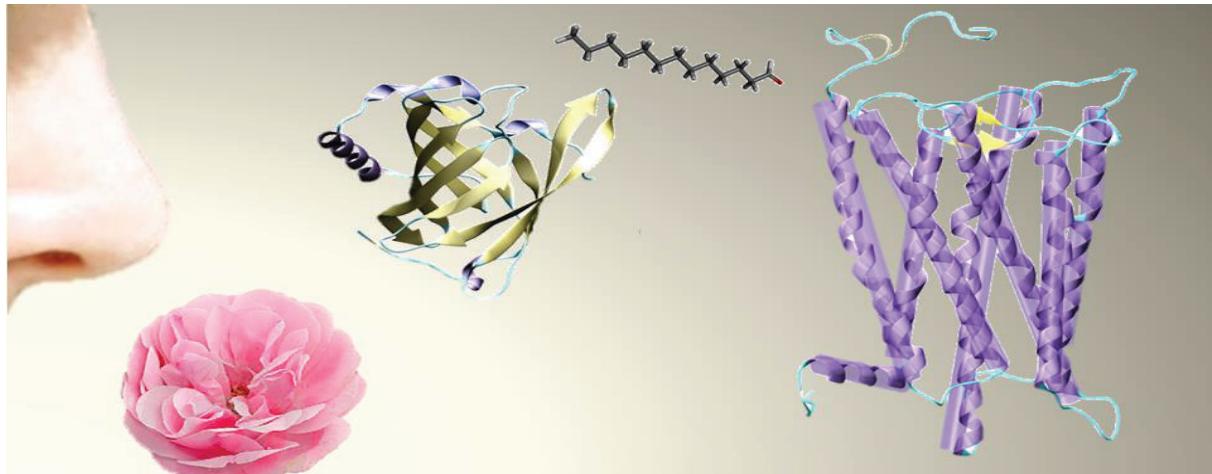
L'utilisation des méthodes théoriques alliées aux validations expérimentales est particulièrement adaptée à l'étude de ces protéines et a permis d'apporter de nombreuses connaissances sur leur mode de fonctionnement. Comment ces protéines lient-elles les molécules odorantes ? Comment se comportent-elles de façon dynamique ? Certains résidus sont-ils plus importants que d'autres ? Nous répertorions dans cette revue les études de modélisation moléculaire réalisées sur les protéines impliquées dans l'olfaction ainsi que les grandes avancées qui en ont découlé. Grâce à ces outils, un microscope computationnel est pointé sur notre perception des odeurs.

Nous noterons que malgré l'apport considérable de ce genre d'approche, certaines questions restent en suspens. Ces deux protéines sont-elles capables d'interagir ? De quelle façon les ROs s'activent lorsque le message chimique transmis par molécule odorante est capté ?

Article 2:

A computational microscope focused on the sense of smell

Claire A. de March & Jérôme Golebiowski*, Biochimie 107 (2014) 3-10



"He knew that the senses, no less than the soul, have their spiritual mysteries to reveal. And so he would now study perfumes"

Oscar Wilde, The picture of Dorian Gray

Keywords: Olfactory Receptor, Odorant-Binding Protein, Molecular modeling, Structure-function

Abstract

In this article, we review studies of the protagonists of the perception of smell focusing on Odorant-Binding Proteins and Olfactory Receptors. We notably put forward studies performed by means of molecular modeling, generally combined with experimental data. Those works clearly emphasize that computational approaches are now a force to reckon with. In the future, it will certainly be more and more used, notably in the framework of a computational microscope meant to observe how the laws of physics govern the biomolecular systems originating our sense of smell.

Introduction

We perceive exogenous odorant chemicals through an extraordinary subtle and sensitive system. The sense of smell belongs to the five commonly admitted senses, already cited by Plato in *Theaetetus*, IV centuries BCE. This sense endows us with the perception of chemicals present in our environment and with identification of their intensity with a very high discriminating power.[1] With the sense of taste, it is however considered as a minor sense with respect to senses of hearing, seeing and touching. As a matter of fact, quite few people know of how to define an individual lacking the perception of smell (anosmia) or taste (ageusia) but everybody knows what is deafness or blindness. Our brain has nonetheless developed a very powerful mechanism for detecting odorants. At the cellular level, the perception of smell is triggered by activation of our Olfactory Receptor Neurons (ORN). As for the molecular level, volatile odorants driven by the inspired air cross our nasal cavity and reach our olfactory mucosa, itself protected by the olfactory mucus that prevents it from drying out (Figure 1).

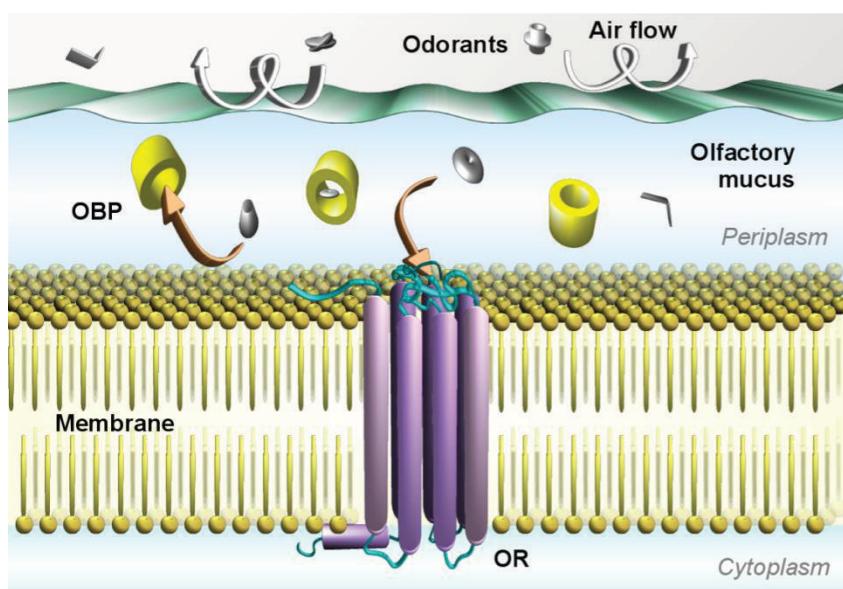


Figure 1. Schematic representation of the molecular protagonists involved in the perception of smell at the Olfactory Receptor Neuron surface. Odorants molecules (grey) cross the air flow to reach the

olfactory mucus (blue). Odorant Binding Proteins (yellow) help solubilize them. Odorants then bind the trans-membrane Olfactory Receptors (purple) to eventually trigger their activation.

This aqueous layer constitutes a physico-chemical barrier for odorant molecules. Indeed the highly hydrophobic character of odorant molecules hampers them from accumulating within this aqueous phase. Within this mucus, the Odorant Binding Proteins (OBP) are the first biological protagonists met by odorants. Despite the fact that their involvement in odor perception is not clearly established, they are thought to bind odorant molecules and transport them to the ORN cilia. At the surface of these ORN cilia, Olfactory Receptors (OR) constitute the cornerstone of the perception of smell. The activation of an OR by an odorant triggers ORN membrane depolarization. These ORNs electric signals spread in our brain and constitute the information we call an odor.

The perception of smell proceeds through a combinatorial code, in which the relation between the odorant space and the receptor space is not a bijection i.e. some odorants can activate several receptors while a single receptor can recognize many diverse odorants.[2]

The virtually infinite number of odorant molecules makes it impossible to experimentally screen their potency with respect to the bimolecular protagonists of smell. As a consequence, computational approaches will be of great help to decipher the role of these proteins. They are sufficiently robust to be considered as a computational microscope.[3] In this mini-review, we propose a survey of the knowledge gained on those proteins using molecular modeling. The first paragraph reports on studies dealing with OBP while the second covers studies of OR.

Odorant Binding Proteins

OBP are small soluble proteins belonging to the family of lipocalin.[4] They are secreted in the nasal mucus at the quite high concentration of ~ 10 mM. The exact function of these proteins is not clearly elucidated. The lipocalin family is notably made up of transport proteins, such as Retinol Binding Protein (RPB), β -lactoglobulin or Bilin Binding proteins (BBP) for example. These proteins share the property of binding hydrophobic ligands to transport them to an *ad hoc* receptor. By similarity, this suggests that OBP help solubilizing odorants within the mucus to finally activate the receptor, as it has been clearly evidenced in the fruit fly for example.[5]

OBP are found in a variety of species including cow, pig, rabbit, mouse, rat, elephant and of course human.[6-8] Different OBP subtypes have been observed in the same species. For example the mouse has four OBP sub-types, the rabbit three and at least eight can be found in the porcupine. In the rat, three OBP have been described with quite different sequences and binding properties.[9-11] OBP molecular weight is between 17 and 22 kDa. Most OBP are observed as monomers, such as

porcine OBP-1, rat OBP-1, rat OBP-3 or human OBP, while some others are found as dimers, such as for the special case of bovine OBP. OBP heterodimers have also been observed in mouse.[12]

Binding properties

OBP reversibly bind odorants with affinities in the micromolar range.[7] As discussed earlier, the fact that a few subtypes are present in most species suggests that OBP do not show a high specificity for a given chemical family. Porcine OBP have been however shown to be post-translationally modified by phosphorylation, generating a diversity of OBP isoforms with specific binding properties.[13]

Although, no preferential binding was put forward for the native porcine and bovine OBP. Studies of the three rat OBP revealed intriguing odorant specificities that fulfil each other to encompass all chemical families. Rat OBP-1 preferentially binds heterocyclic compounds such as pyrazine derivatives while OBP-2 appears to be more specific for long-chain aliphatic aldehydes and carboxylic acids. Rat OBP-3 was described as associated to odorants composed of saturated or unsaturated ring structure.[9, 14-16] Human OBP-IIA appears unspecific at first sight, as it can associate to various odorant types with dissociation constants in the micromolar range.[6] A chemical specificity for aldehydes, either aliphatic or aromatic was however revealed.[17]

Sequence & structure

All lipocalins have generally a low percentage of identity (~25%). The maximal known identity of 42% is found between a monomer of the bovine and the porcine OBP and the lowest concern rat OBP-2 (12-19%). The hallmarks of their sequence are scarce: a GxW motif is found ~ 15-20 residues away from the N-terminus, two cysteine residues in the middle of the sequence and a glycine residue at the C-terminal end.

From a structural point of view, OBP share the typical lipocalin fold. They are composed of an eight-strand β -barrel, flanked by seven loops. An α -helix followed by a small β -sheet is present at the C-terminal end, as shown in figure 2. A cysteine residue on the D strand of the β -barrel is engaged in a disulfide bridge with the C-terminal domain.

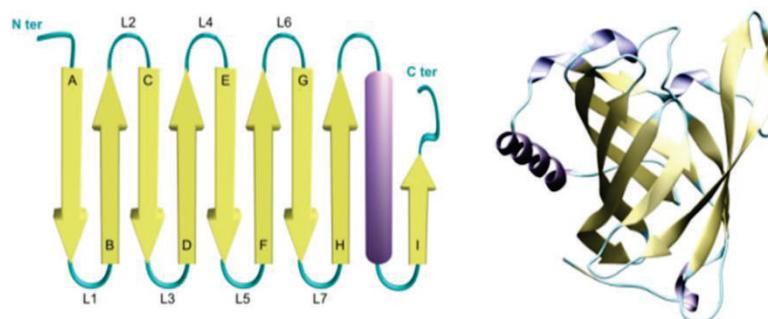


Figure 2. Secondary structure and typical 3 dimensional structure of a mammalian Odorant-Binding Protein. β -sheets are shown in yellow, α -helices in magenta and loops in cyan.

The β -barrel constitutes a *calyx* bearing a *lipophilic* cavity. This property is at the origin of their family name (lipo/calyx). The bovine OBP is a bit special. It is made up of an oligomer, where two lipocalin structures have undergone a so-called ‘domain swapping’ through an exchange of their α -helices. [18]

Molecular modeling of OBP

The rather easy way to express those hydrosoluble proteins has allowed solving crystal structures for many of them. It follows that at the exception of a few cases, molecular modeling studies are most of the time based-on an experimental structure. In many research articles, molecular modeling was mainly used with the purpose of identifying the protein binding site. It’s only more recently that the function of these proteins was investigated by means of theoretical approaches.

Through a series of explicit solvent molecular dynamics studies, the dynamic behavior of OBP has been described. The volume of the binding cavity was computed to $\sim 500 \text{ \AA}^3$ and showed fluctuations between 400 and 800 \AA^3 during the MD simulations. Although occluded from the bulk water, the binding cavity showed transient openings, notably at the junction between loop1 and strands D and E of the β -barrel.[19] Unconstrained[20] or constrained[21] molecular dynamics simulations observed the binding or unbinding of odorants through the opening of this part of the protein, confirming that it is the main access from the bulk to the binding cavity. Gratifyingly, the crucial role of a highly conserved tyrosine residue (Y82) initially pointed out in the simulations was experimentally confirmed afterwards.[13]

Once bound to the OBP, the odorant is engaged in an opportunistic interaction with the protein involving a few hydrogen-bonds.[22] If hydrogen bonds are observed, they possess short residence time since the position of the odorant within the cavity is subjected to hydrophobic contacts. The number of these contact together with the size of the odorant appear to be the main factor originating high affinity.[16, 19] More elaborated protocols allowed computing the affinities between these unspecific proteins and series of odorants. State-of-the-art free energies were able to predict with a high accuracy the differential affinity between closely related ligands, showing that a computational deorphanization of OBP is possible.[23] More modest models, solely based on ligands structures, such as pharmacophore approaches, were proved unable to predict the affinity while docking and end-point MM-GBSA[24] free energy approaches perform equally at reproducing an experimental ranking between odorants with respect to their affinity.[25]

Focusing on human OBP, Tatchoff et al. proposed a model to investigate the intriguing selectivity of hOBP-IIA for aldehydes and acids.[17] This protein has no crystal structure available and a homology-based model was built using the human tear lipocalin as a template. Although the binding

mode is not associated to strong hydrogen bonds between the ligand and the cavity residues,[19, 22] the authors identified three lysine residues within the binding site, putatively involved in hydrogen bonding with odorants. Lys 112 was identified as strongly involved in aldehydes recognition. An implicit solvent molecular dynamics study has been performed on the same protein with the strongly bound citral and undecanal.[26] The experimental higher affinity for undecanal is qualitatively recovered by a MM-GBSA approach and is associated to a stronger hydrogen-bond with Lys 112. More generally, the larger affinity of hOBP-IIA for aldehydes is proposed to result from an equilibrium involving a Schiff base between the amine side-chain of Lys 112 and the aldehyde function.

Other simulations were run in parallel with experimental studies with the purpose of using OBP as biosensors.[27] Mutants of the dimeric bovine OBP at the domain swapping produces monomeric OBP, which is then strongly similar to that of the pork. Based on molecular dynamics studies combined to infra-red spectroscopy, ionic and hydrophobic interactions were shown to be mainly responsible for the protein stability in both the porcine and the bovine OBP.[28-30] Similar results were found for the rat OBP3 where a molten-globule state was identified just prior to denaturation.[31] All these denaturations or modifications of the protein are reversible, but the ligand binding cannot occur anymore.[32] In addition, the dimeric wild-type bovine OBP was thermally more stable than the mutant monomer.[29] Generally, these structures are stable at very high pressure[33] or very low pH.[34] The presence of a ligand notably helps stabilizing intra-molecular interactions within the protein, as observed in molecular dynamics simulations performed at 2000 bars.[33]

Olfactory Receptors

Olfactory Receptors can be regarded as the cornerstone of the perception of smell. OR genes are a multigenic family corresponding to more than 2% of our genome. They have been discovered by Buck and Axel in 1991.[35] In humans, men and women show equivalent OR as the genes coding for these proteins are spread on every chromosomes at the exception of chromosomes 8, 20 and Y.[36, 37] The number of genes coding for human OR reaches ~1000, of which a part may be tagged as pseudo-genes, leading to 396 functional OR.[38] Differences among modern humans, Neanderthals and Denisovans was recently put forward.[39] Interindividual differential expression of OR among humans was recently put forward. Anyway a conserved set of 90 OR was detected among a set of 26 individuals. In addition, OR expression was shown to be independent of age, sex or smoking status but the level of expression of some genes vary with age.[40] Rats and dogs appear as particularly well equipped for discrimination between odorants as they possess 872 and 1201 functional OR, respectively.[41] Counter-intuitively, our smelling performance is not far removed from those of the

latter as our cognitive power compensates the smaller number of functional OR.[42] Notice also that some pseudo-genes can nevertheless lead to a functional OR as shown for hOR1E3P,[43] for which molecular modeling suggested a non-typical class A GPCR structure.[44]

Glusman et al. proposed a classification in which OR belonging to a family or a sub-family possess an analogy of at least 40% and 60%, respectively.[45] For example, OR1A1 and OR1A2 show at least 60% of sequence homology while OR1A1 and OR1G1 possess at least 40% of sequence homology.

Interestingly, OR are expressed in other organs like heart,[46] male germinal cells,[47] spleen and pancreas.[48] See ref[49] for a review on those ectopic expressions.

At the cellular level, OR are expressed by Olfactory Receptor Neurons (ORN). Each ORN expresses only one type of OR.[50-52] As a consequence, the study of the response of an OR is equivalent to the study of the response of a neuron making a direct connection between molecular sciences and neuroscience. Malnic et al. shown that one odorant molecule is able to activate several OR and one OR can be activated by several odorants. The perception of smell results from a combinatorial code arising from different OR.[2] The fascinating complexity of odor prediction based on the structure of an odorant has already been discussed.[53, 54]

Binding properties

Raming et al. assessed for the first time in 1993 that OR were able to bind odorant molecules.[55] OR are activated by odorants with affinities associated to micromolar to nanomolar *in vitro* EC50.[56] Intriguingly, these potencies are comparable to the affinities found for an odorant with an OBP.[7]

Many studies contributed to the understanding of binding properties of OR by experimental deorphanization data, either on mouse[57-69] or human[43, 56, 70-86]. Generally, OR can be split into two main families. So called broadly tuned OR are activated by odorants having different chemical properties and also varying in size, shape and even stereochemistry. This is typically the case for the human OR2W1[56] or mouse OR256-17[69]. Other OR exhibit a narrow range of odorant recognition, such as OR7D4 in humans,[84] specifically activated by androstenone and androstadienone, or OR256-8 in mouse, primarily responding to linear aliphatic aldehydes, alcohols and esters.[69]

The large variability of OR sequences endows our brain with the potential detection of all chemicals responding to an odorant criteria *viz.* a moderate molecular weight ($<400 \text{ g}\cdot\text{mol}^{-1}$), a high lipophilicity, a weak polarity and a certain water solubility. Even molecules containing inorganic atoms like silicon instead of carbon are still able to elicit a response of an OR.[87]

Sequence & structure

OR belong to the class-A of the family of G-Protein Coupled Receptors (GPCR). As such, they all share some sequence and structural features: seven trans-membrane helices (TM1 to TM7) connected by extra-cellular and intra-cellular loops as depicted in figure 3. A short eighth helix is present in the intracellular part. As for the sequence, the main conserved motifs of this family, considered as the hallmark of OR are as follows:

- GN in TM1
- LHxPMYFFLxxLSxxD in TM2
- MAYDRYVAICxPLxY in TM3
- SY in TM5
- KAFSTCxSH in TM6
- PxLNPxIYSLNR in TM7

Two cysteines residues are highly conserved in OR, one at the top of TM3 and one in the extra-cellular loop 2, likely to make a disulfide bridge as observed in experimental structures of GPCR.[88]

The sequence identity between OR and other class-A GPCR is less than 20%. [89]

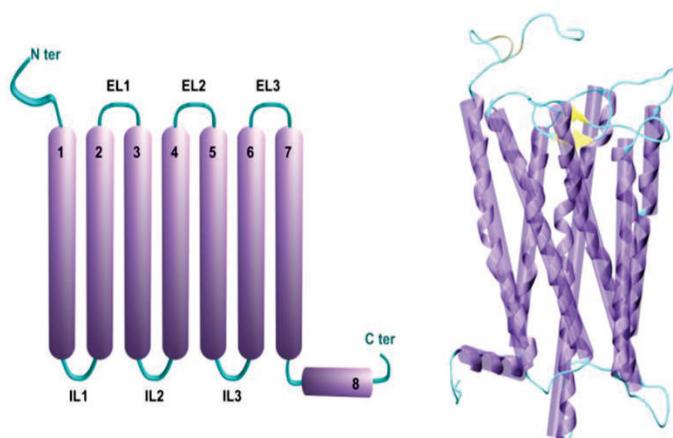


Figure 3. Secondary structure and modeled 3 dimensional structure of a mammalian Olfactory Receptor. α -helices are shown in magenta and loops in cyan.

The seven trans-membrane helices form a bundle in which a cavity is dedicated to bind odorants. The eighth helix is located at the intracellular membrane surface and is supposed to act like an anchoring point to the membrane.[90]

Molecular modeling of OR

No experimental structure of any OR is actually available. As a consequence, molecular modeling offers a unique opportunity to catch a glimpse into the atomic level details of those objects. With the recent advances in force field development together with increase of computational power and the

availability of pertinent templates, molecular simulation of OR is becoming widely used. Two main philosophies can be invoked for building an OR i) so-called *ab initio* methods consider first principles approaches where the whole structure can be predicted by optimizing the interaction energy between sub parts of it,[91] ii) homology based methods consist in getting inspired from available known structures belonging to the same family as the target.[92] Applied to OR, *ab initio* protocols use classical laws of physics to construct a 3D structure on the basis of the amino-acid sequence. First, the trans-membrane domains are predicted using hydrophobic analysis combined with some sequence alignments. The helix bundle is then built by optimizing relative positions of the helices with respect to each other. The rotational orientation of the helices in a lipid bilayer is then energy minimized using a force field approach. The loops are finally built to obtain a final structure of the OR.[93] This has the advantage of getting rid of experimental templates but it is anyway subjected to the quality of the force field used for both the protein and the lipids. The method has to date been successfully applied to several mouse, rat and human OR.[93-98] In a pioneering work, this method was applied to the mouse OR S25. The receptor was built and embedded in a solvated DPPC membrane prior to molecular dynamics simulations. The OR binding site was described at the atomic level and the two experimentally determined agonists hexanol and heptanol were recovered as strong binders.[94] The rat OR-I7 was used to assess the accuracy of molecular modeling. Short simulations of this OR in a vacuum proposed a model to identify the ligands binding pathway.[98] In another study based on the analysis of the binding site residues, *in silico* site-directed mutagenesis revealed that Lys 164^{4.57} was strongly involved in electrostatic contact with the bond agonists. The super script 4.57 refers to the Ballesteros-Weinstein numbering.[99] The most conserved residue in each TM domain was attributed following the alignment published by us.[88] The phenotype of V206I^{5.33} was however hardly explained since the model predicted that this residue points towards the membrane.[97] Despite their 95% sequence identity, the subtlety between the two rat and mouse orthologs of OR-I7 lead to crucial structural differences as shown by the model of Hall et al. Differences in helical bends, tilts and translations of the transmembrane helices were put forward and sufficiently change the residues lining the binding pocket to modify the response of odorants.[96]

Generally, the binding sites are predicted to be located between TM3, TM5 and TM6, as expected from an analysis of variable residues in OR sequences.[89, 100] It was further confirmed by several studies based on homology modeling and site-directed mutagenesis.[63, 72, 101] A comparison between homology modeling and this *ab initio* method was briefly discussed by Charlier et al. The two models show very similar structures, with a Root Mean Square Deviation of 6.8 Å for the whole bundle structure and 1.6 Å when only considering residues lining the binding cavity.[80]

Mid way between *ab initio* and homology modeling, some OR have been built using the low resolution structure of bovine rhodopsin, further refined by first principles calculations as shown in the pioneering work on rat OR5 bound to lyral OR-I7.[97, 98, 102] In this protocol, the TM sequences are predicted from a hydrophobicity profile and the structure of each helix is obtained through an energy minimization. The OR bundle is finally built under the constraint of the 7.5 Å resolution rhodopsin template. Alternatively, the use of a homology protocol can be considered as a starting point for building the TM domains prior to a refinement based on optimisation of hydrophobic packing.[103, 104]

The performance of such approaches notwithstanding, strict homology modeling is the major source of 3D models. The quality of homology-based structures will now benefit from the increase of available experimental structures of class A GPCR.[105-121] Once the latter is chosen, the main difficulty stands in an accurate sequence alignment between the target OR and the template.[88] In most studies, OR were built based on either the rhodopsin or the β2-adrenergic folds.[63, 72, 74, 75, 80, 87, 101, 122, 123] Opsin was recently proposed as a good structural model for OR.[120]

Generally, molecular modeling reveals extremely powerful when used to guide site-directed mutagenesis experiments. Such combined approaches help understanding the nature of the binding sites of OR and how some odorants can be recognized with a high affinity. The mouse Eugenol OR (mOR-EG) is a typical example. The binding site has been identified as made up of mainly hydrophobic residues in contact with some odorants (notably V109^{3.37}, F206^{5.37}, L212^{5.43}, F252^{6.57}, I256^{6.61}, L259^{6.64}).[63, 101] The gained knowledge about these residues allows identifying some chemical features required to build potent agonists. While the rhodopsin derived model of Katada et al. proposes that agonists should have three substitutions on a benzene ring,[63] the β2-adrenergic inspired model of Baud et al. suggests a higher variability in the structure.[101] Anyway, despite the use of different templates, both studies recover similar binding sites. This suggests, as also observed for hOR1G1,[80, 123] that the impact of the template is deemed rather minor when focusing on residues of the binding cavities.

As for the odorant-OR interaction, a detailed *in silico* analysis of the binding cavities properties between the two paralogs hOR1A1 and hOR1A2 has identified the few amino-acids responsible for the subtle differential affinity between (S)-(-)-citronellol and (S)-(-)-citronellal.[75] Similarly, the accuracy of these computational approaches was spectacularly illustrated in a study on hOR2AG1 combining molecular modeling and site-directed mutagenesis. The authors managed to reprogram the OR through the identification of some residues governing the recognition spectrum. The single F206V mutation made the receptor highly sensitive to isoamylbenzoate.[72]

Considering ligand affinities, molecular modeling of the hOR1D2 bound to liliol, bourgeonal and their silicon analogues emphasized similar positions within the binding site but differential binding strength.[87] Generally the affinity between odorants and the receptor is computed by means of a docking score (see for a typical example the study comparing the mouse OR 912-93 and its human ortholog.[93] In the framework of virtual screening using a database of more than 500 odorants, docking scores have been used for the purpose of deorphanization of the mouse OR42-3. The performance of this approach was close to 50%.[124] Some studies however reported more elaborated protocols to compare the experimental OR activation with an accurate free energy of binding, taking into account flexibility through molecular dynamics simulations (see figure 4). Nevertheless, these simulations are subjected to limitations associated to state-of-the-art molecular modeling. At best, only structural insights can nowadays be obtained, with eventually an estimation of the affinity between the ligand and the receptor. Notice that in the case of GPCR, ligand potency is not fully correlated with affinity. Typically, agonists and antagonists both show high affinities but opposite potencies. Their discrimination will require sampling of the OR activation, which has to date not been performed. Such large scale conformational movements have only been observed for a few GPCR which structures were experimentally known.[125-128]

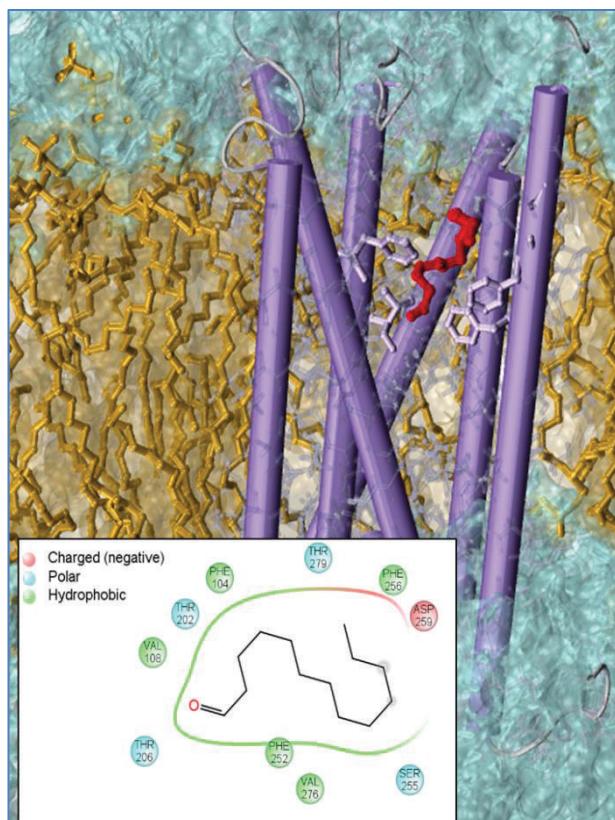


Figure 4. State-of-the-art molecular modeling explicitly takes into account the atomic level details of both the complex between an odorant (tridecanal in red) and a receptor (hOR1G1 in purple) and its physiological environment made up of a phospholipid bilayer (yellow) solvated by water (blue).

Typically, this system contains around 50,000 atoms and is subjected to a molecular dynamics simulation.[88] Inset: chemical structure of the ligand and identification of amino-acids in contact with the latter.

Agonists and non-agonists were discriminated for hOR1G1 using molecular dynamics simulations further analyzed with the MM-GBSA protocol. The decomposition of the free energy of binding on a per residue basis revealed a multimodal way of binding where the cavity residues are differentially involved depending on the odorant chemical family.[80, 129, 130] A comparison between analogs of helional was performed on hOR3D1 using the thermodynamic integration method. It also allowed splitting odorants between binders and non-binders by comparing their computed affinities with the receptor.[74] In the special case of rat OR-I7 bound to aldehydes, a hybrid Quantum Mechanics / Molecular Mechanics (QM/MM) methodology was set up to investigate the formation of a Schiff base between the odorants and a lysine residue directed into the cavity, as suggested in a previous work.[96] The free energy of binding trend within the series recovered experiment.[122]

Although it is widely accepted that heterogeneous local lipid environment can affect structural, dynamics or functional properties of membrane receptors,[131] these medium-induced phenomena have not been studied so far on OR. This is probably because long scale molecular dynamics simulations are only emerging in the area of the description of the perception of smell.

Interaction between OBP and OR

OBP and OR are the first molecular protagonists of the perception of smell. Odorant molecules are in a first step solubilized within the olfactory mucus by OBP. In a second step, interaction with the OR triggers neuron's response. The link between these two proteins is unresolved in human olfaction but some clues suggest synergistic effects. In fact, experimental studies reported that OBP and OR are interacting together.[132-135] The interaction between the OBP and the receptor would preserve OR function at high odorant concentration. The bell-shaped dose-response curve observed for human OR 17-40 activated by helional was turned into an S-shaped curve in presence of rat OBP-1F. On the basis of this observation, the authors put forward a model where an OBP binds a dimer of OR but atomic level information about this contact and the way an odorant could be transferred from the OBP to the OR remains to be uncovered.[132] This will probably be a topic of high interest that will involve the community of molecular modeling in next few years.

Conclusion

The sense of smell is triggered at the molecular level by the interaction between chemicals called odorants with several biological protagonists. Among them, Odorant-Binding Proteins and Olfactory Receptors are well identified as being involved in the response of the Olfactory Receptor Neurons

expressed in our Olfactory Bulb. Here, we have reviewed some characteristics of these fascinating molecular machines and notably focused on how molecular modeling provides detailed atomic insights into the way these proteins behave or bind odorants. Such a tool can describe the structure, predict the affinity with odorant or alternatively guide experiment to gain knowledge on their function. Although all models focused on a single biomolecular system, a study focused on the interaction between OBP and OR would be of high interest since few experimental studies put forward a modulation of the OR response in the presence of an OBP.[132-135]

More generally, with the evolution of computational technologies and force field development, molecular dynamics of OR can now be performed on the millisecond time scale, which is *a priori* sufficient to observe the receptor activation.[125, 126, 128, 136] Such atomic level models could represent the rise of a bottom-up computational neuroscience. In these emerging methods, one could predict an Olfactory Receptor Neuron response solely based-on the chemical formula of an odorant and the sequence of the gene coding for the OR housed by this neuron. These researches would be highly beneficial for neuroscientists that try deciphering the neural code associated to the perception of smell. The industry of flavors and fragrances will also benefit from these researches by being able to rationally design an odorant on the basis of its combinatorial interaction with the receptors, as it is commonly used in the pharmacological domain.

Acknowledgments

The referees provided helpful comments. This work was supported by a grant from the foundation Edmond Roudnitska, under the aegis of Foundation de France to Claire A. de March. Claire A. de March thanks Giract for a PhD bursary. This work is part of the Olfactome project, funded by the région Provence Alpes Côte d'Azur.

References

- [1] C. Bushdid, M.O. Magnasco, L.B. Vosshall, A. Keller, Humans Can Discriminate More than 1 Trillion Olfactory Stimuli, *Science* 343 (2014) 1370-1372.
- [2] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial Receptor Codes for Odors, *Cell* 96 (1999) 713-723.
- [3] E.H. Lee, J. Hsin, M. Sotomayor, G. Comellas, K. Schulten, Discovery through the computational microscope, *Structure* 17 (2009) 1295-1306.
- [4] D.R. Flower, Multiple molecular recognition properties of the lipocalin protein family, *J. Mol. Recognit.* 8 (1995) 185-195.
- [5] P. Xu, R. Atkinson, D.N. Jones, D.P. Smith, *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons, *Neuron* 45 (2005) 193-200.
- [6] L. Briand, C. Eloit, C. Nespolous, V. Bezirard, J.C. Huet, C. Henry, F. Blon, D. Trotier, J.C. Pernollet, Evidence of an odorant-binding protein in the human olfactory mucus: location, structural characterization, and odorant-binding properties, *Biochemistry* 41 (2002) 7241-7252.
- [7] P. Pelosi, The role of perireceptor events in vertebrate olfaction, *Cellular and Molecular Life Sciences* 58 (2001) 503-509.
- [8] M. Tegoni, P. Pelosi, F. Vincent, S. Spinelli, V. Campanacci, S. Grolli, R. Ramoni, C. Cambillau, Mammalian odorant binding proteins, *Biochimica et biophysica acta* 1482 (2000) 229-240.
- [9] D. Lobel, M. Jacob, M. Volkner, H. Breer, Odorants of different chemical classes interact with distinct odorant binding protein subtypes, *Chemical senses* 27 (2002) 39-44.
- [10] D. Lobel, S. Marchese, J. Krieger, P. Pelosi, H. Breer, Subtypes of odorant-binding proteins, *European journal of biochemistry* 254 (1998) 318-324.
- [11] D. Lobel, J. Strotmann, M. Jacob, H. Breer, Identification of a third rat odorant-binding protein (OBP3), *Chemical senses* 26 (2001) 673-680.
- [12] D. Pes, M. Mameli, I. Andreini, J. Krieger, M. Weber, H. Breer, P. Pelosi, Cloning and expression of odorant-binding proteins Ia and Ib from mouse nasal tissue, *Gene* 212 (1998) 49-55.
- [13] P. Nagnan-Le Meillour, C. Le Danvic, F. Brimau, P. Chemineau, J.C. Michalski, Phosphorylation of native porcine olfactory binding proteins, *Journal of chemical ecology* 35 (2009) 752-760.
- [14] A.J. Borysik, L. Briand, A.J. Taylor, D.J. Scott, Rapid odorant release in mammalian odour binding proteins facilitates their temporal coupling to odorant signals, *Journal of molecular biology* 404 (2010) 372-380.
- [15] L. Briand, C. Nespolous, V. Perez, J.J. Remy, J.C. Huet, J.C. Pernollet, Ligand-binding properties and structural characterization of a novel rat odorant-binding protein variant, *European journal of biochemistry* 267 (2000) 3079-3089.
- [16] C. Nespolous, L. Briand, M.M. Delage, V. Tran, J.C. Pernollet, Odorant binding and conformational changes of a rat odorant-binding protein, *Chemical senses* 29 (2004) 189-198.
- [17] L. Tatchoff, C. Nespolous, J.C. Pernollet, L. Briand, A single lysyl residue defines the binding specificity of a human odorant-binding protein for aldehydes, *FEBS Lett* 580 (2006) 2102-2108.
- [18] M. Tegoni, R. Ramoni, E. Bignetti, S. Spinelli, C. Cambillau, Domain swapping creates a third putative combining site in bovine odorant binding protein dimer, *Nature structural biology* 3 (1996) 863-867.
- [19] J. Golebiowski, S. Antonczak, D. Cabrol-Bass, Molecular dynamics studies of odorant binding protein free of ligand and complexed to pyrazine and octenol, *THEOCHEM Journal of Molecular Structure* 763 (2006) 165-174.
- [20] J. Golebiowski, S. Antonczak, S. Fiorucci, D. Cabrol-Bass, Mechanistic events underlying odorant binding protein chemoreception, *Proteins* 67 (2007) 448-458.

- [21] E. Hajjar, D. Perahia, H. Debat, C. Nespolous, C.H. Robert, Odorant binding and conformational dynamics in the odorant-binding protein, *Journal of Biological Chemistry* 281 (2006) 29929-29937.
- [22] F. Vincent, S. Spinelli, R. Ramoni, S. Grolli, P. Pelosi, C. Cambillau, M. Tegoni, Complexes of porcine odorant binding protein with odorant molecules belonging to different chemical classes, *Journal of molecular biology* 300 (2000) 127-139.
- [23] L. Charlier, C. Nespolous, S. Fiorucci, S. Antonczak, J. Golebiowski, Binding free energy prediction in strongly hydrophobic biomolecular systems, *Physical Chemistry Chemical Physics* 9 (2007) 5761-5771.
- [24] I. Massova, P. Kollman, Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding, *Perspectives in Drug Discovery and Design* 18 (2000) 113-135.
- [25] J. Golebiowski, J. Topin, L. Charlier, L. Briand, Interaction between odorants and proteins involved in the perception of smell: the case of odorant-binding proteins probed by molecular modelling and biophysical data, *Flavour and Fragrance Journal* 27 (2012) 445-453.
- [26] L. Charlier, D. Cabrol-Bass, J. Golebiowski, How does human odorant binding protein bind odorants? The case of aldehydes studied by molecular dynamics, *C. R. Chim.* 12 (2009) 905-910.
- [27] O.V. Stepanenko, O.V. Stepanenko, M. Staiano, I.M. Kuznetsova, K.K. Turoverov, S. D'Auria, The Quaternary Structure of the Recombinant Bovine Odorant-Binding Protein Is Modulated by Chemical Denaturants, *OBPS ONE* 9 (2014) e85169.
- [28] O.V. Stepanenko, A. Marabotti, I.M. Kuznetsova, K.K. Turoverov, C. Fini, A. Varriale, M. Staiano, M. Rossi, S. D'Auria, Hydrophobic interactions and ionic networks play an important role in thermal stability and denaturation mechanism of the porcine odorant-binding protein, *Proteins: Structure, Function & Bioinformatics* 71 (2008) 35-44.
- [29] A. Marabotti, A. Scirè, M. Staiano, R. Crescenzo, V. Aurilia, F. Tanfani, S. D'Auria, Wild-Type and Mutant Bovine Odorant-Binding Proteins To Probe the Role of the Quaternary Structure Organization in the Protein Thermal Stability, *Journal of Proteome Research* 7 (2008) 5221-5229.
- [30] M. Staiano, S. D'Auria, A. Varriale, M. Rossi, A. Marabotti, C. Fini, O.V. Stepanenko, I.M. Kuznetsova, K.K. Turoverov, Stability and Dynamics of the Porcine Odorant-Binding Protein, *Biochemistry* 46 (2007) 11120-11127.
- [31] A. Scire, A. Marabotti, M. Staiano, L. Briand, A. Varriale, E. Bertoli, F. Tanfani, S. D'Auria, Structure and stability of a rat odorant-binding protein: another brick in the wall, *Journal of Proteome Research* 8 (2009) 4005-4013.
- [32] S. Marchal, A. Marabotti, M. Staiano, A. Varriale, T. Domaschke, R. Lange, S. D'Auria, Under Pressure That Splits a Family in Two. The Case of Lipocalin Family, *OBPS ONE* 7 (2012) e50489.
- [33] M. Staiano, M. Saviano, P. Herman, Z. Grycznyski, C. Fini, A. Varriale, A. Parracino, A.B. Kold, M. Rossi, S. D'Auria, Time-resolved fluorescence spectroscopy and molecular dynamics simulations point out the effects of pressure on the stability and dynamics of the porcine odorant-binding protein, *Biopolymers* 89 (2008) 284-291.
- [34] A. Mazzini, E. Polverini, M. Parisi, R.T. Sorbi, R. Favilla, Dissociation and unfolding of bovine odorant binding protein at acidic pH, *Journal of Structural Biology* 159 (2007) 82-91.
- [35] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [36] S. Rouquier, S. Taviaux, B.J. Trask, V. Brand-Arpon, G.v.d. Engh, J. Demaille, D. Giorgi, Distribution of olfactory receptor genes in the human genome, *Nature Genetics* 18 (1998) 243-250.
- [37] B. Malnic, P.A. Godfrey, L.B. Buck, The human olfactory receptor gene family, *Proceedings of the National Academy of Sciences, USA* 101 (2004) 2584-2589.

- [38] A. Matsui, Y. Go, Y. Niimura, Degeneration of Olfactory Receptor Gene Repertoires in Primates: No Direct Link to Full Trichromatic Vision, *Molecular Biology and Evolution* 27 (2010) 1192-1200.
- [39] G.M. Hughes, E.C. Teeling, D.G. Higgins, Loss of Olfactory Receptor Function in Hominin Evolution, *OBPS ONE* 9 (2014) e84714.
- [40] C. Verbeurgt, F. Wilkin, M. Tarabichi, F. Gregoire, J.E. Dumont, P. Chatelain, Profiling of Olfactory Receptor Gene Expression in Whole Human Olfactory Mucosa, *OBPS ONE* 9 (2014) e96333.
- [41] P. Quignon, M. Giraud, M. Rimbault, P. Lavigne, S. Tacher, E. Morin, E. Retout, A.-S. Valin, K. Lindblad-Toh, J. Nicolas *et al*, The dog and rat olfactory receptor repertoires, *Genome Biology* 6 (2005) R83.
- [42] G.M. Shepherd, The Human Sense of Smell: Are We Better Than We Think?, *OBPS Biology* 2 (2004) e146.
- [43] V. Matarazzo, O. Clot-Faybesse, B. Marcet, G. Guiraudie-Capraz, B. Atanasova, G. Devauchelle, M. Cerutti, P. Etiévant, C. Ronin, Functional Characterization of Two Human Olfactory Receptors Expressed in the Baculovirus Sf9 Insect Cell System, *Chemical senses* 30 (2005) 195-207.
- [44] P. Lai, G. Bahl, M. Gremigni, V. Matarazzo, O. Clot-Faybesse, C. Ronin, C. Crasto, An olfactory receptor pseudogene whose function emerged in humans: a case study in the evolution of structure–function in GPCRs, *J Struct Funct Genomics* 9 (2008) 29-40.
- [45] G. Glusman, A. Bahar, D. Sharon, Y. Pilpel, J. White, D. Lancet, The olfactory receptor gene superfamily: data mining, classification, and nomenclature, *Mammalian Genome* 11 (2000) 1016-1023.
- [46] G. Drutel, J.M. Arrang, J. Diaz, C. Wisnewsky, K. Schwartz, J.C. Schwartz, Cloning of OL1, a putative olfactory receptor and its expression in the developing rat heart, *Receptors and Channels* 3 (1995) 33-40.
- [47] M. Parmentier, F. Libert, S. Schurmans, S. Schiffmann, A. Lefort, D. Eggerickx, C. Ledent, C. Mollereau, C. Gerard, J. Perret *et al*, Expression of members of the putative olfactory receptor gene family in mammalian germ cells, *Nature* 355 (1992) 453-455.
- [48] P. Blache, L. Gros, G. Salazar, D. Bataille, Cloning and Tissue Distribution of a New Rat Olfactory Receptor-like (OL2), *Biochemical and Biophysical Research Communications* 242 (1998) 669-672.
- [49] N. Kang, J. Koo, Olfactory receptors in non-chemosensory tissues, *BMB reports* 45 (2012) 612-622.
- [50] P. Mombaerts, Odorant receptor gene choice in olfactory sensory neurons: the one receptor–one neuron hypothesis revisited, *Current Opinion in Neurobiology* 14 (2004) 31-36.
- [51] S. Serizawa, K. Miyamichi, H. Sakano, One neuron–one receptor rule in the mouse olfactory system, *Trends in Genetics* 20 (2004) 648-653.
- [52] S.H. Fuss, A. Ray, Mechanisms of odorant receptor gene choice in *Drosophila* and vertebrates, *Molecular and Cellular Neuroscience* 41 (2009) 101-112.
- [53] C.S. Sell, On the Unpredictability of Odor, *Angewandte Chemie International Edition* 45 (2006) 6254-6261.
- [54] A. Triller, E.A. Boulden, A. Churchill, H. Hatt, J. Englund, M. Spehr, C.S. Sell, Odorant–Receptor Interactions and Odor Percept: A Chemical Perspective, *Chemistry & Biodiversity* 5 (2008) 862-886.
- [55] K. Raming, J. Krieger, J. Strotmann, I. Boekhoff, S. Kubick, C. Baumstark, H. Breer, Cloning and expression of odorant receptors, *Nature* 361 (1993) 353-356.
- [56] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, J.D. Mainland, Odor Coding by Mammalian Receptor Repertoire, *Science Signaling* 2 (2009) ra9.
- [57] I. Boekhoff, K. Touhara, S. Danner, J. Ingles, M.J. Lohse, H. Breer, R.J. Lefkowitz, Phosducin, Potential Role in Modulation of Olfactory Signaling, *Journal of Biological Chemistry* 272 (1997) 4606-4612.

- [58] K. Touhara, S. Sengoku, K. Inaki, A. Tsuboi, J. Hirono, T. Sato, H. Sakano, T. Haga, Functional identification and reconstitution of an odorant receptor in single olfactory neurons, *Proceedings of the National Academy of Sciences* 96 (1999) 4040-4045.
- [59] R.C. Araneda, A.D. Kini, S. Firestein, The molecular receptive range of an odorant receptor, *Nature Neuroscience* 3 (2000) 1248-1255.
- [60] K. Kajiya, K. Inaki, M. Tanaka, T. Haga, H. Kataoka, K. Touhara, Molecular Bases of Odor Discrimination: Reconstitution of Olfactory Receptors that Recognize Overlapping Sets of Odorants, *Journal of Neuroscience* 21 (2001) 6018-6025.
- [61] T. Bozza, P. Feinstein, C. Zheng, P. Mombaerts, Odorant Receptor Expression Defines Functional Units in the Mouse Olfactory System, *Journal of Neuroscience* 22 (2002) 3033-3043.
- [62] S. Katada, T. Nakagawa, H. Kataoka, K. Touhara, Odorant response assays for a heterologously expressed olfactory receptor, *Biochemical and Biophysical Research Communications* 305 (2003) 964-969.
- [63] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural Basis for a Broad But Selective Ligand Spectrum of a Mouse Olfactory Receptor: Mapping the Odorant-Binding Site, *Journal of Neuroscience* 25 (2005) 1806-1815.
- [64] Y. Oka, M. Omura, H. Kataoka, K. Touhara, Olfactory receptor antagonism between odorants, *The EMBO Journal* 23 (2004) 120-126.
- [65] Y. Oka, S. Katada, M. Omura, M. Suwa, Y. Yoshihara, K. Touhara, Odorant Receptor Map in the Mouse Olfactory Bulb: In Vivo Sensitivity and Specificity of Receptor-Defined Glomeruli, *Neuron* 52 (2006) 857-869.
- [66] T. Abaffy, H. Matsunami, C.W. Luetje, Functional analysis of a mammalian odorant receptor subfamily, *Journal of Neurochemistry* 97 (2006) 1506-1518.
- [67] X. Grosmaire, S.H. Fuss, A.C. Lee, K.A. Adipietro, H. Matsunami, P. Mombaerts, M. Ma, SR1, a Mouse Odorant Receptor with an Unusually Broad Response Profile, *Journal of Neuroscience* 29 (2009) 14545-14552.
- [68] A. Kato, K. Touhara, Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization, *Cellular and Molecular Life Sciences* 66 (2009) 3743-3753.
- [69] J. Li, R. Haddad, S. Chen, V. Santos, C.W. Luetje, A broadly tuned mouse odorant receptor that detects nitrotoluenes, *Journal of Neurochemistry* 121 (2012) 881-890.
- [70] G. Sanz, C. Schlegel, J.-C. Pernollet, L. Briand, Comparison of Odorant Specificity of Two Human Olfactory Receptors from Different Phylogenetic Classes and Evidence for Antagonism, *Chemical senses* 30 (2005) 69-80.
- [71] G. Sanz, T. Thomas-Danguin, H. Hamdani el, C. Le Poupon, L. Briand, J. Pernollet, E. Guichard, A. Tromelin, Relationships between molecular structure and perceived odor quality of ligands for a human olfactory receptor, *Chemical senses* 33 (2008) 639-653.
- [72] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angewandte Chemie International Edition* 51 (2012) 1274-1278.
- [73] T. Dahoun, L. Grasso, H. Vogel, H. Pick, Recombinant Expression and Functional Characterization of Mouse Olfactory Receptor mOR256-17 in Mammalian Cells, *Biochemistry* 50 (2011) 7228-7235.
- [74] C. Anselmi, A. Buonocore, M. Centini, R.M. Facino, H. Hatt, The human olfactory receptor 17-40: Requisites for fitting into the binding pocket, *Computational Biology and Chemistry* 35 (2011) 159-168.
- [75] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *Journal of Structural Biology* 159 (2007) 400-412.
- [76] M. Spehr, G. Gisselmann, A. Poplawski, J.A. Riffell, C.H. Wetzel, R.K. Zimmer, H. Hatt, Identification of a Testicular Odorant Receptor Mediating Human Sperm Chemotaxis, *Science* 299 (2003) 2054-2058.

- [77] B.L. Cook, D. Steuerwald, L. Kaiser, J. Graveland-Bikker, M. Vanberghem, A.P. Berke, K. Herlihy, H. Pick, H. Vogel, S. Zhang, Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4, *Proceedings of the National Academy of Sciences* 106 (2009) 11925-11930.
- [78] P. Olsson, M. Laska, Human Male Superiority in Olfactory Sensitivity to the Sperm Attractant Odorant Bourgeonal, *Chemical senses* 35 (2010) 427-432.
- [79] J.F. McRae, J.D. Mainland, S.R. Jaeger, K.A. Adipietro, H. Matsunami, R.D. Newcomb, Genetic Variation in the Odorant Receptor OR2J3 Is Associated with the Ability to Detect the "Grassy" Smelling Odor, *cis*-3-hexen-1-ol, *Chemical senses* 37 (2012) 585-593.
- [80] L. Charlier, J. Topin, C. Ronin, S.-K. Kim, W. Goddard, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cellular and Molecular Life Sciences* (2012) 1-9.
- [81] K.A. Adipietro, J.D. Mainland, H. Matsunami, Functional Evolution of Mammalian Odorant Receptors, *OBPS Genetics* 8 (2012) e1002821.
- [82] V. Jacquier, H. Pick, H. Vogel, Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds, *Journal of Neurochemistry* 97 (2006) 537-544.
- [83] E.M. Neuhaus, A. Mashukova, W. Zhang, J. Barbour, H. Hatt, A Specific Heat Shock Protein Enhances the Expression of Mammalian Olfactory Receptor Proteins, *Chemical senses* 31 (2006) 445-452.
- [84] A. Keller, H.Y. Zhuang, Q.Y. Chi, L.B. Vosshall, H. Matsunami, Genetic variation in a human odorant receptor alters odour perception, *Nature* 449 (2007) 468-U466.
- [85] I. Menashe, T. Abaffy, Y. Hasin, S. Goshen, V. Yahalom, C.W. Luetje, D. Lancet, Genetic Elucidation of Human Hyperosmia to Isovaleric Acid, *OBPS Biology* 5 (2007) e284.
- [86] J.D. Mainland, A. Keller, Y.R. Li, T. Zhou, C. Trimmer, L.L. Snyder, A.H. Moberly, K.A. Adipietro, W.L.L. Liu, H. Zhuang *et al*, The missense of smell: functional variability in the human odorant receptor repertoire, *Nature Neuroscience* 17 (2014) 114-120.
- [87] L. Doszczak, P. Kraft, H.-P. Weber, R. Bertermann, A. Triller, H. Hatt, R. Tacke, Prediction of Perception: Probing the hOR17-4 Olfactory Receptor Model with Silicon Analogues of Bourgeonal and Lilial, *Angewandte Chemie International Edition* 46 (2007) 3367-3371.
- [88] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, in: C.J. Crasto (Eds.), *Olfactory Receptors, Molecular Modelling of Odorant/Olfactory Receptor Complexes*, 2013, pp. 53-65.
- [89] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Science* 13 (2004) 240-254.
- [90] V.A. Avlani, K.J. Gregory, C.J. Morton, M.W. Parker, P.M. Sexton, A. Christopoulos, Critical Role for the Second Extracellular Loop in the Binding of Both Orthosteric and Allosteric G Protein-coupled Receptor Ligands, *Journal of Biological Chemistry* 282 (2007) 25677-25686.
- [91] N. Vaidehi, W.B. Floriano, R. Trabanino, S.E. Hall, P. Freddolino, E.J. Choi, G. Zamanakos, W.A. Goddard, Prediction of structure and function of G protein-coupled receptors, *Proceedings of the National Academy of Sciences* 99 (2002) 12622-12627.
- [92] E. Krieger, S.B. Nabuurs, G. Vriend, in: P.E. Bourne, H. Weissig (Eds.), *Structural Bioinformatics, Homology Modeling*, 2005, pp. 509-523.
- [93] P. Hummel, N. Vaidehi, W.B. Floriano, S.E. Hall, W.A. Goddard, Test of the Binding Threshold Hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93, *Protein Science* 14 (2005) 703-710.
- [94] W.B. Floriano, N. Vaidehi, W.A. Goddard, M.S. Singer, G.M. Shepherd, Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor, *Proceedings of the National Academy of Sciences* 97 (2000) 10712-10716.
- [95] W.B. Floriano, N. Vaidehi, W.A. Goddard, Making Sense of Olfaction through Predictions of the 3-D Structure and Function of Olfactory Receptors, *Chemical senses* 29 (2004) 269-290.

- [96] S.E. Hall, W.B. Floriano, N. Vaidehi, W.A. Goddard, Predicted 3-D Structures for Mouse I7 and Rat I7 Olfactory Receptors and Comparison of Predicted Odor Recognition Profiles with Experiment, *Chemical senses* 29 (2004) 595-616.
- [97] M.S. Singer, Analysis of the Molecular Basis for Octanal Interactions in the Expressed Rat I7 Olfactory Receptor, *Chemical senses* 25 (2000) 155-165.
- [98] P.C. Lai, M.S. Singer, C.J. Crasto, Structural Activation Pathways from Dynamic Olfactory Receptor–Odorant Interactions, *Chemical senses* 30 (2005) 781-792.
- [99] J.A. Ballesteros, H. Weinstein, Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, *Methods Neurosci.* 25 (1995) 366-428.
- [100] Y. Pilpel, D. Lancet, The variable and conserved interfaces of modeled olfactory receptor proteins, *Protein Science* 8 (1999) 969-977.
- [101] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The Mouse Eugenol Odorant Receptor: Structural and Functional Plasticity of a Broadly Tuned Odorant Binding Pocket, *Biochemistry* 50 (2011) 843-853.
- [102] M. Afshar, R.E. Hubbard, J. Demaille, Towards structural models of molecular recognition in olfactory receptors, *Biochimie* 80 (1998) 129-135.
- [103] P.C. Lai, C.J. Crasto, Beyond Modeling: All-Atom Olfactory Receptor Model Simulations, *Frontiers in Genetics* 3 (2012)
- [104] P.C. Lai, B. Guida, J. Shi, C.J. Crasto, Preferential Binding of an Odor Within Olfactory Receptors: A Precursor to Receptor Activation, *Chemical senses* 39 (2014) 107-123.
- [105] T. Okada, M. Sugihara, A.-N. Bondar, M. Elstner, P. Entel, V. Buss, The Retinal Conformation and its Environment in Rhodopsin in Light of a New 2.2Å Crystal Structure, *Journal of molecular biology* 342 (2004) 571-583.
- [106] S.G.F. Rasmussen, H.-J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R.P. Ratnala, R. Sanishvili, R.F. Fischetti *et al*, Crystal structure of the human β 2 adrenergic G-protein-coupled receptor, *Nature* 450 (2007) 383-387.
- [107] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.-J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka *et al*, High-Resolution Crystal Structure of an Engineered Human β 2-Adrenergic G Protein–Coupled Receptor, *Science* 318 (2007) 1258-1265.
- [108] R.O. Dror, D.H. Arlow, D.W. Borhani, M. \emptyset . Jensen, S. Piana, D.E. Shaw, Identification of two distinct inactive conformations of the β 2-adrenergic receptor reconciles structural and biochemical observations, *Proceedings of the National Academy of Sciences* 106 (2009) 4689-4694.
- [109] R. Horst, J.J. Liu, R.C. Stevens, K. Wuthrich, β (2)-Adrenergic Receptor Activation by Agonists Studied with F-19 NMR Spectroscopy, *Angew. Chem.-Int. Edit.* 52 (2013) 10762-10765.
- [110] A.M. Ring, A. Manglik, A.C. Kruse, M.D. Enos, W.I. Weis, K.C. Garcia, B.K. Kobilka, Adrenaline-activated structure of β (2)-adrenoceptor stabilized by an engineered nanobody, *Nature* 502 (2013) 575-579.
- [111] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G.W. Leslie, C.G. Tate, G.F.X. Schertler, Structure of a β 1-adrenergic G-protein-coupled receptor, *Nature* 454 (2008) 486-491.
- [112] E.Y.T. Chien, W. Liu, Q. Zhao, V. Katritch, G. Won Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov *et al*, Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist, *Science* 330 (2010) 1091-1095.
- [113] B. Wu, E.Y.T. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi *et al*, Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, *Science* 330 (2010) 1066-1071.
- [114] S. Granier, A. Manglik, A.C. Kruse, T.S. Kobilka, F.S. Thian, W.I. Weis, B.K. Kobilka, Structure of the δ -opioid receptor bound to naltrindole, *Nature* 485 (2012) 400-404.

- [115] J.F. White, N. Noinaj, Y. Shibata, J. Love, B. Kloss, F. Xu, J. Gvozdenovic-Jeremic, P. Shah, J. Shiloach, C.G. Tate *et al*, Structure of the agonist-bound neurotensin receptor, *Nature* 490 (2012) 508-513.
- [116] A.A. Thompson, W. Liu, E. Chun, V. Katritch, H. Wu, E. Vardy, X.-P. Huang, C. Trapella, R. Guerrini, G. Calo *et al*, Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic, *Nature* 485 (2012) 395-399.
- [117] A. Manglik, A.C. Kruse, T.S. Kobilka, F.S. Thian, J.M. Mathiesen, R.K. Sunahara, L. Pardo, W.I. Weis, B.K. Kobilka, S. Granier, Crystal structure of the μ -opioid receptor bound to a morphinan antagonist, *Nature* 485 (2012) 321-326.
- [118] L. Zhu, Q. Zhao, B. Wu, Structure-based studies of chemokine receptors, *Current Opinion in Structural Biology* 23 (2013) 539-546.
- [119] A.K. Shukla, A. Manglik, A.C. Kruse, K. Xiao, R.I. Reis, W.-C. Tseng, D.P. Staus, D. Hilger, S. Uysal, L.-Y. Huang *et al*, Structure of active β -arrestin-1 bound to a G-protein-coupled receptor phosphopeptide, *Nature* 497 (2013) 137-141.
- [120] J.H. Park, T. Morizumi, Y. Li, J.E. Hong, E.F. Pai, K.P. Hofmann, H.-W. Choe, O.P. Ernst, Opsin, a Structural Model for Olfactory Receptors?, *Angewandte Chemie International Edition* 52 (2013) 11021-11024.
- [121] H. Wu, D. Wacker, M. Mileni, V. Katritch, G.W. Han, E. Vardy, W. Liu, A.A. Thompson, X.-P. Huang, F.I. Carroll *et al*, Structure of the human κ -opioid receptor in complex with JDTic, *Nature* 485 (2012) 327-332.
- [122] M.D. Kurland, M.B. Newcomer, Z. Peterlin, K. Ryan, S. Firestein, V.S. Batista, Discrimination of Saturated Aldehydes by the Rat I7 Olfactory Receptor, *Biochemistry* 49 (2010) 6302-6304.
- [123] G. Launay, S. Teletchea, F. Wade, E. Pajot-Augy, J.F. Gibrat, G. Sanz, Automatic modeling of mammalian olfactory receptors and docking of odorants, *Protein Engineering Design and Selection* 25 (2012) 377-386.
- [124] S. Bavan, B. Sherman, C.W. Luetje, T. Abaffy, Discovery of Novel Ligands for Mouse Olfactory Receptor MOR42-3 Using an *In Silico* Screening Approach and *In Vitro* Validation, *OBPS ONE* 9 (2014) e92064.
- [125] R.O. Dror, D.H. Arlow, P. Maragakis, T.J. Mildorf, A.C. Pan, H. Xu, D.W. Borhani, D.E. Shaw, Activation mechanism of the β 2-adrenergic receptor, *Proceedings of the National Academy of Sciences* 108 (2011) 18684-18689.
- [126] R. Nygaard, Y. Zou, Ron O. Dror, Thomas J. Mildorf, Daniel H. Arlow, A. Manglik, Albert C. Pan, Corey W. Liu, Juan J. Fung, Michael P. Bokoch *et al*, The Dynamic Process of β 2-Adrenergic Receptor Activation, *Cell* 152 (2013) 532-542.
- [127] Y. Miao, S.E. Nichols, J.A. McCammon, Free energy landscape of G-protein coupled receptors, exOBPPred by accelerated molecular dynamics, *Physical Chemistry Chemical Physics* (2014) 6398-6406.
- [128] E.N. Laricheva, K. Arora, J.L. Knight, C.L. Brooks, Deconstructing Activation Events in Rhodopsin, *Journal of the American Chemical Society* 135 (2013) 10906-10909.
- [129] J. Golebiowski, L. Charlier, J. Topin, S. Fiorucci, S. Antonczak, in: V. Ferreira, R. Lopez (Eds.), Flavour Science, Chapter 96 - Molecular Features Underlying the Chemoreception of Odorant Binding Proteins and Olfactory Receptors. Insights from Molecular Modeling and Biophysical Data, San Diego, 2014, pp. 519-523.
- [130] J. Topin, C.A. de March, L. Charlier, C. Ronin, S. Antonczak, J. Golebiowski, Discrimination between Olfactory Receptor agonists and non-agonists, *Chemistry A Eur. J.* 20 (2014) 10227-10230.
- [131] J. Zimmerberg, K. Gawrisch, The physical chemistry of biological membranes, *Nature Chemical Biology* 2 (2006) 564-567.
- [132] J. Vidic, J. Grosclaude, R. Monnerie, M.-A. Persuy, K. Badonnel, C. Baly, M. Caillol, L. Briand, R. Salesse, E. Pajot-Augy, On a chip demonstration of a functional role for odorant binding protein in the preservation of olfactory receptor activity at high odorant concentration, *Lab on a Chip* 8 (2008) 678-688.

- [133] H.J. Ko, T.H. Park, Enhancement of odorant detection sensitivity by the expression of odorant-binding protein, Biosensors and Bioelectronics 23 (2008) 1017-1023.
- [134] H. Ko, S. Lee, E. Oh, T. Park, Specificity of odorant-binding proteins: a factor influencing the sensitivity of olfactory receptor-based biosensors, Bioprocess Biosyst Eng 33 (2010) 55-62.
- [135] V. Matarazzo, N. Zsürger, J.-C. Guillemot, O. Clot-Faybesse, J.-M. Botto, C.D. Farra, M. Crowe, J. Demaille, J.-P. Vincent, J. Mazella *et al*, Porcine Odorant-binding Protein Selectively Binds to a Human Olfactory Receptor, Chemical Senses 27 (2002) 691-701.
- [136] L.C.T. Pierce, R. Salomon-Ferrer, C. Augusto F. de Oliveira, J.A. McCammon, R.C. Walker, Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics, Journal of Chemical Theory and Computation 8 (2012) 2997-3002.

Modélisation moléculaire des récepteurs olfactifs

Articles 3 et 4 - La modélisation moléculaire des ROs : de la séquence à la structure.

L'étude des relations structure-odeur sur une base physiologiquement inspirée nécessite la prise en compte des récepteurs olfactifs. La modélisation moléculaire permet notamment de prédire la structure d'une protéine ou de son complexe avec une molécule, d'en évaluer l'énergie d'interaction et d'observer son comportement dynamique. Les protocoles permettant de réaliser chacune de ces étapes sont variés.

La modélisation moléculaire des complexes ROs/odorants

Les RCPGs sont couramment étudiés par des méthodes théoriques. Afin de se rapprocher des conditions physiologiques du récepteur, les modèles étudiés sont souvent constitués d'une membrane solvatée dans laquelle le récepteur est imbriqué. Les lois de la physique sont appliquées aux atomes formant ces systèmes pour en observer le comportement dynamique. Ces protocoles ont été employés avec succès sur des RCPGs non-olfactifs et sont utilisés dans le cadre de nos études sur les ROs. Dans l'article « Molecular modelling of odorant / Olfactory Receptor complexes » (article 3), la méthodologie couramment associée à l'étude des ROs est décrite en détails.

L'alignement de séquences : le point crucial

La reconstruction par homologie d'une protéine repose essentiellement sur l'alignement de la séquence de la protéine ciblée avec celles des modèles de référence. Les séquences des ROs de mammifères partagent des motifs d'acides aminés conservés. Ces derniers caractérisent les protéines de cette sous-famille. Ils sont présents dans chacune des sept hélices des ROs et représentent un point d'ancrage pour aligner de façon non-ambigüe leur séquence. Aucune structure expérimentale de RO n'est disponible. Toutefois, la structure de 21 RCPGs non-olfactifs de classe A a été élucidée expérimentalement. Ils sont, à ce jour, le modèle de référence optimal pour la reconstruction de la structure des ROs par homologie. La principale difficulté est donc d'aligner les motifs de ROs à ceux des RCPGs non-olfactifs. Ces deux types de récepteurs partagent une grande partie de leurs séquences mais sont néanmoins dissemblables sur quelques points. Comment pallier le manque d'homologie dans certaines parties des séquences pour réaliser l'alignement pertinent ?

La mutagenèse dirigée permet d'identifier expérimentalement le rôle d'un résidu ciblé. Elle consiste à remplacer ce résidu par un autre à travers la modification du gène de la protéine. Le mutant ainsi obtenu est exprimé dans la cellule et sa réponse à un composé ou plusieurs est quantifiée. Des indices sur le rôle de cet acide aminé sont obtenus grâce à l'éventuelle modification du phénotype de la protéine. Sous la contrainte de ces données, les acides aminés sont alignés non pas par leur nature

Articles 3-4

mais par leur fonction dans le récepteur. La pertinence de l'alignement obtenu permet de construire des modèles de ROs validés par des contraintes expérimentales, dans la limite des connaissances actuelles et dans l'attente qu'une structure expérimentale d'un RO soit disponible.

L'alignement de séquence est un point crucial pour la reconstruction par homologie d'un RO, mais plusieurs questions restent ouvertes. Quelle est l'influence du modèle de référence choisi ? Le protocole de reconstruction par homologie est-il le plus adapté ? Les structures obtenues par différents protocoles sont-elles comparables ?

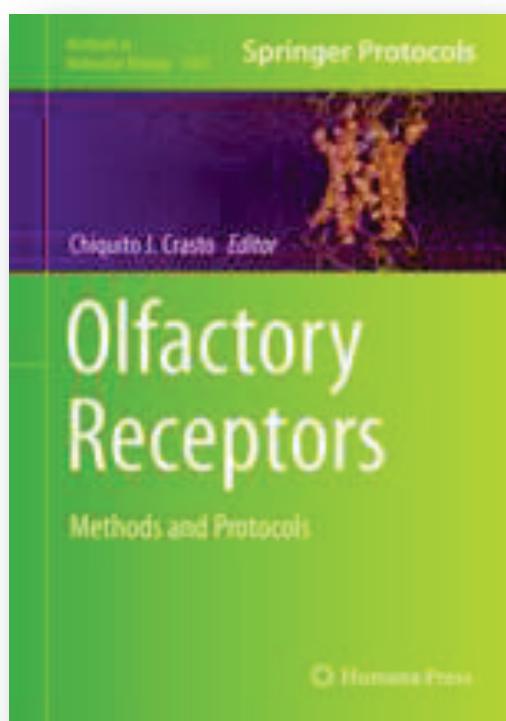
Dans cette partie, les protocoles de modélisation moléculaire appliqués à l'étude des ROs sont décrits. J'ai donc conçu l'alignement proposé dans l'article de méthodologie « Molecular modelling of odorant / Olfactory Receptor complexes » (article 3). J'ai développé les détails de la conception de cet alignement, discuté des choix des modèles et comparé les différentes approches de modélisation dans l'article « G Protein-Coupled Odorant Receptors: from sequence to structure » (article 4). Cet article a été réalisé avec la collaboration de nos homologues du Caltech, le Dr. Soo-Kyung Kim et le Pr. William Goddard III.

Article 3:

Molecular modelling of odorant / Olfactory Receptor complexes

Landry Charlier, Jérémie Topin, Claire A. de March, Chiquito J. Crasto, Jérôme Golebiowski

In Olfactory Receptors - Methods in Molecular Biology, volume 1003, 2013, pp 53-65



Keywords: Molecular modelling, homology, 3D-structure, docking, GPCR, olfactory receptor, MM-GBSA.

Abstract

Providing a rationale that associates a chemical structure of an odorant to its induced perception has been sought for a long time. To achieve this, a detailed atomic structure of both the odorant and the olfactory receptor must be known. State-of-the-art techniques to model the 3D structure of an olfactory receptor in complex with various odorants are presented here. These range from sequence alignment with known structures, to molecular dynamics simulations in a realistic environment

1. Introduction

The connection between the structure of an odorant and its role in the perception of odors has been long sought. At the molecular level, the perception of smell is rooted in the activation of olfactory neurons, each of them housing an olfactory receptor (see [1] for a review). The large number of olfactory receptors (ORs) gives rise to the idea of olfaction being associated with a combinatorial signal[2] which is, for now, virtually impossible to mechanistically elucidate in terms of the odorant's chemical structure. Prior to the discovery of olfactory receptors, Structure-Property relationship studies of odors projected promising results.[3] Unfortunately, they suffered from a major limitation since the role of the structure of the ORs was not considered.

With the recent advances in both computing power and bioinformatics methodologies, molecular modeling has evolved as a force that allows us a glimpse into the nature of OR-odor interactions at a molecular level. Although one of the major quests of molecular modeling, *i.e.* the description of a protein structure on the basis of its amino acid sequence is within reach,[4] ORs structural description represents a particularly challenging task, since these proteins belong to the family of G-protein-Coupled-Receptors (GPCRs)- membrane protein, whose structures are notoriously difficult to structurally characterize.

Two distinct approaches, *ab initio* or homology modeling methods, can be used to overcome the lack of X-ray structures of ORs. *Ab initio* protocols use classical laws of physics to construct a 3D structure on the basis of the amino acid sequence (see [5] for an example on GPCRs). Homology modeling methods aim at predicting the structure of a protein of interest from a set of experimentally known structures. The latter is the major methodology used to build theoretical 3D models. Here, we describe the materials and methods to build a full, atomistic 3D structure of an OR, both free or complexed with an odorant. Additionally, embedding the OR or its complex with an odorant within a solvated phospholipid bilayer and system relaxation by molecular dynamics simulations is presented. The present protocol has notably been used in ref.[6]

2. Materials

2.1. Sequence comparison and alignment

1. The protein sequences can be downloaded on servers like the Protein Information Resource (PIR: <http://pir.georgetown.edu/>) or the Human Olfactory Receptor Data Explorer (HORDE: <http://genome.weizmann.ac.il/horde/>).
2. The alignment can be carried out by freeware such as Jalview (<http://www.jalview.org/>) or directly on servers like PIR.

2.2. 3D structure building

1. Modeller (a homology or comparative modeling of protein three-dimensional structures freeware)
2. Any 3D-visualization software (VMD, Chimera, etc.)

2.3. Ligand docking

1. Any docking software (AUTODOCK VINA, GOLD, etc.)
2. Files containing the 3D structure of the odorants (in *pdb* or *mol2* file format)

2.4. Membrane embedding

1. GROMACS molecular modeling software
2. A file containing the 3D structure of the phospholipid membrane (DPPC or POPC for example can be found on P. Tieleman's Web page:
http://moose.bio.ucalgary.ca/index.php?page=Structures_and_Topologies).
3. InflateGro script to build the membrane environment around the OR
(<http://moose.bio.ucalgary.ca/files/inflategro>).
4. Alternatively, the Maestro from Schrödinger, inc. (<http://schrodinger.com>). The software has a membrane building protocol.

2.5. Molecular dynamics

1. Any molecular dynamics software (AMBER, GROMACS, NAMD, etc.).
2. Choose a force field developed for both proteins and lipids.

3. Methods

3.1. Sequence alignment

The alignment should be performed with at least one known experimental structure of a class A GPCR. Currently, the X-ray structures of at least height different GPCRs have been solved, the bovine

Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65
rhodopsin,[7] the human adenosine A2a receptor,[8] the turkey beta-1 adrenergic,[9] the human beta-2 adrenergic receptor,[10] the human CXCR4 chemokine receptor,[11] the human dopamine D3 receptor,[12] the human histamine receptor H1,[13] and the S1P1 sphingosine 1-phosphate receptor.[14] The sequence identity between these proteins and the ORs is less than 20% and thus, should ideally be compensated by the use of experimental data. In spite of this low sequence identity, several regions are nonetheless conserved in GPCRs, such as the *GN* residues in helix 1, the *DRY* segment in helix 3. A cysteine residue present in helix 3 and another one in the extracellular loop 2 (EL2) form a conserved cysteine bridge, observed in experimental structures. A second cysteine pair is conserved in 98% of ORs sequences. Here it involves the cysteine residues 169 and 189. Moreover, other residues are generally conserved within the OR family:[15]

- *LHXP MYFFL* in the beginning of helix 2,
- *MAYDRYVVA/CXPLXY* in the end of helix 3,
- *SY* in helix 5,
- *KAFSTCXSH* in helix 6
- *PMLNPF/YSLRN* in helix 7.

1. Paste both the OR and experimentally known GPCR sequences in FASTA format in the PIR Web site (<http://pir.georgetown.edu/pirwww/search/multialn.shtml>). Several sequences are particularly useful for ORs studies:

- (a) The multiple alignment published by Man *et al*[16] aligned five ORs with five others GPCRs (including the rhodopsin sequence). It helps to correctly align conserved regions.
- (b) The sequences of the other experimental GPCR structures (notably the human β 1 and β 2 adrenergic receptors).
- (c) Other studies have been performed on ORs and can complete the set.[6, 17-25] (see **Note 1**)

2. Retrieve the alignment and open it in the Jalview software. Eventually, manually assess the alignment to ensure that the conserved regions discussed above are correctly aligned. Figure 1 shows an alignment performed to produce the structure of the human OR, *OR1G1* (also named hOR17-209).

PARTIE 2 : Modélisation moléculaire des récepteurs olfactifs

Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65

Figure 1. Alignment of OR1G1, several ORs and other class A GPCR sequences. The predicted transmembrane helices are shown as gray bars. For a better reading, the first residues of mOR-480, as well as the central part of the third intracellular loop of hD2 (human dopamine D2 receptor), ha2A (human 2A adrenergic receptor), A2AA (human 2A adrenergic receptor) were omitted. Some residues are identified as: O amino acid residues common to OR, A amino acid residues common to class A GPCR, G amino acid residues allowing the GPCR activation, C amino acid residues including in a potential OR conformational change, P OR amino acid residue in contact with the G protein.

3.2. From sequence to 3D structure

1. On the Protein Data Bank Web site, obtain the 3D structure of: the bovine Rhodopsin (PDB id: 1U19), the human β 2 adrenergic receptor (PDB id: 2RH1), the turkey β 1 adrenergic receptor (PDB id: 2VT4) and the human A2a adenosine receptor (PDB id: 3EML). Modify each PDB file by removing water molecules, β -factor, etc., keeping only the receptor's residues and the natural ligand.
 2. Prepare the MODELLER input file. Specify the cysteine residues forming S-S bridges. Note that two sulphur bridges are highly probable in ORs based on the alignment with known class A GPCRs

Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65

such as β 2-Adrenergic receptor. In our case these two bridges involve cysteine residue 97 and 179 and also 169 and 189. Be sure to consider that the ligand (retinal in 1U19, cyanopindolol in 2VT4, carazolol in 2RH1 or ZM241385 in 3EML) located within the reference protein is inserted in the OR during the building process. It will build the OR structure with a quite large internal cavity, in which the odorants will later be docked.

3. Generate a large number of putative structures (say 50-100) to allow maximum flexibility among these structures. The template structure considered for the building procedure can be either one of the solved X-ray structures of GPCR (rhodopsin, β 1, β 2, or adenosine receptor, CXCR4 chemokine receptor, dopamine D3 receptor, histamine receptor, lipid GPCR, muscarinic M2 receptor, nociception receptor, delta opioid receptor, kappa opioid receptor, mu opioid receptor) or one obtained by using all as templates for the OR target. Typically, currently, most OR models have been built with rhodopsin as a template. One of the major differences between the crystallized GPCRs lies in the conformation of the extracellular loop 2 (EL2). This EL2 was shown to be important for ligand recognition in class A GPCRs[26] and its structure has to be considered with care. The several families of structures will be analysed in the next steps. (see **Note 2**)

3.3. 3D model analysis and validation

The analysis and validation procedure depends on both classical physicochemical considerations (hydrophobicity, hydrophilicity, etc.) and on more elaborated protocols, such as the Ramachandran plots, which represents the amino acids conformations.

1. View each model using the visualization software.
2. Eliminate the apparently badly folded structures (for example, those with entwined loops).
3. Eliminate models where too many hydrophobic residues are in the extracellular loops.
4. Check the residues which are in a sphere of five Angstrom units around the ligand of the reference protein and compare with already studied ORs.[6, 17-25] Eliminate models for which either none or too few correspondences are found with ORs built with experimental constraints (site-directed mutagenesis). (see **Note 3**)
5. Check the Ramachandran plot for each remaining model (see Fig. 2). They can be built on Web servers like Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) or on Visual Molecular Dynamics. Select models with the least residues in the outlier regions.

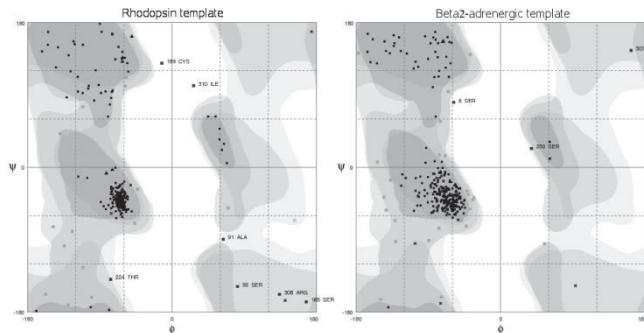


Fig. 2 Ramachandran plot for two homology models based on rhodopsin or β 2-adrenergic templates.

6. One or two structures should be chosen for each family, as shown in Fig. 3.

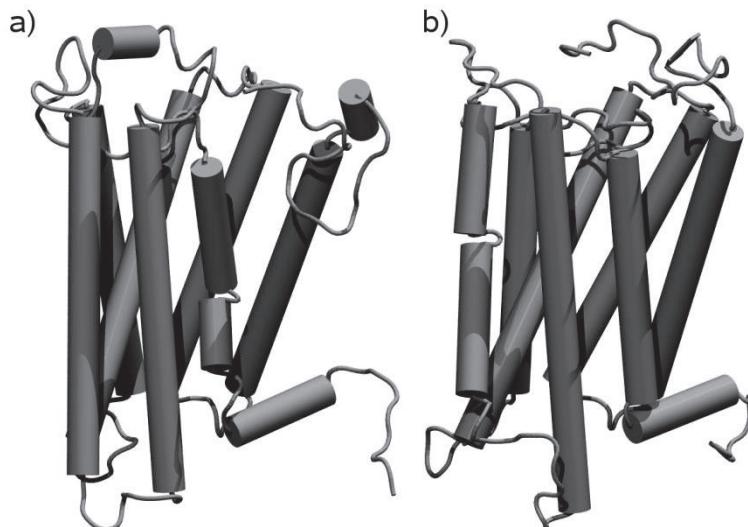


Figure 3. OR1G1 structures built from the β 2-adrenergic (PDB id: 2RH1, a) and rhodopsin (PDB id: 1U19, b) templates.

Figure 3 highlights small deviations between the models built from either the β 2-adrenergic receptor or the rhodopsin structures. These differences arise from various kinks or shifts in the secondary structure of some helices. Moreover, variations in the sequence of the extracellular domain among the GPCRs contribute to the diversity of secondary structures. In this case, however, the residues that constitute the binding cavity are similar, rationalizing the use of these models for ligand-receptor interaction analysis purposes.

3.4. Building the complexes

1. At this step, the ligand of the template protein is always present in the binding site. Remove it manually to enable the docking of the models of odorant molecules.

2. Dock the ligand into the binding site, either with a docking protocol or manually if sufficient evidence exists of a preferred docking configuration (see **Note 4**).
3. Check the environment of the odorant by comparing the ligand-receptor interactions with those described in studies involving site-directed mutagenesis or other pure *in silico* studies. Select the docking conformations involving residues found to be important in site-directed mutagenesis experiments.
4. This stage can be considered the final step, resulting in a 3D-structure of an OR bound to an odorant. Much information is already present in this model, such as the binding affinity estimation, the nature of residues forming the main contacts with the odorants and the orientation of the odorant within the binding site, as shown in Fig. 4. We describe nonetheless, additional and more technical procedure, to refine these data. The relaxation of the structure is likely to slightly change conclusions drawn from this first model. Indeed, the docking protocol has proposed several positions of the ligand within the binding site. Generally, the scoring functions lead to very weak energy differences between the different poses. This suggests that at least two or three conformations may be considered for further refinement. A rescoring function, based on statistics accrued during Molecular Dynamics (MD) simulations should help to get a more accurate model.



Figure 4. Close-up view of the binding site of OR1G1 bound to camphor.

3.5. Membrane embedding

To relax the system in a realistic environment, the phospholipidic membrane, as well as the intra and extracellular medium should be modelled. These steps are rather technical since they require a good knowledge of the use of molecular simulation software.

No odorant should be present in the binding site, since the force-field does not necessarily recognize the odorant atoms. The whole embedding procedure can be done with the unbound OR. The odorant can be reintroduced after OR is stabilized in the membrane. (see **Note 5**). Steps 1 to 7 describe a complex protocol using GROMACS. Step 8 is much simpler alternative using Maestro.

PARTIE 2 : Modélisation moléculaire des récepteurs olfactifs

Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65

1. Check the width of the initial membrane you have built or downloaded. Eventually duplicate it with the *genconf* command of GROMACS in directions specified by the option *-nbox*.
2. Open both the membrane and the OR files using visualization software. Check the position of the membrane with respect to the residues belonging to the OR (hydrophobic residues in the membrane, etc.). A tryptophan residue is for example a good indicator of the membrane position. Its pyrrole functional group can form hydrogen bonds with the polar heads of the lipids while its aromatic cycle interacts with the hydrophobic part of the lipids. If the membrane is badly positioned use the *editconf* command of Gromacs with the option *-center* specifying the coordinates of the centre of the membrane. Use the option *-rotate* to rotate the protein around x, y or z and check that the OR principal axis of inertia is orthogonal to the water/lipid interface.
3. Paste the coordinates of the membrane PDB file in the PDB file of the rotated OR. At this stage, you have probably created a structure where many phospholipids have steric clashes with the receptor. Further refinement is warranted.
4. Generate the topology (*top*) and the coordinates (*gro*) files of the system OR/membrane. Use the *pdb2gmx* command of Gromacs with the *-ignh* command (ignore hydrogen). Choose the force field (for example “*Gromos 54a7*”). During the process, the cysteine bridges are recognized and created. Nonetheless, check that the appropriate cysteine residues are taken into account (if it’s not the case, it indicates that the alignment prior to modeling is incorrect).
5. Inflate the lipids to eliminate steric clashes. Use the *inflategro* script. This script requires a scaling factor and a cut-off, to scale the phospholipids coordinates and deletes those that are too close to the receptor.
6. Slowly deflate the lipids.
 - (a) Create a minimization input file restraining the coordinates of the protein and the phosphor atom of the lipids in the three dimensions and in the z-dimension respectively.
 - (b) Begin minimizing (with Gromacs command *grompp* and *mdrun*).
 - (c) Use the *inflategro* script.
 - (d) Repeat the above steps until the area per lipid containing in the *areaperlipid.dat* file reaches the experimental values (see **Note 6**).[27] Fig. 5 illustrates the different steps of the *inflategro* protocol.

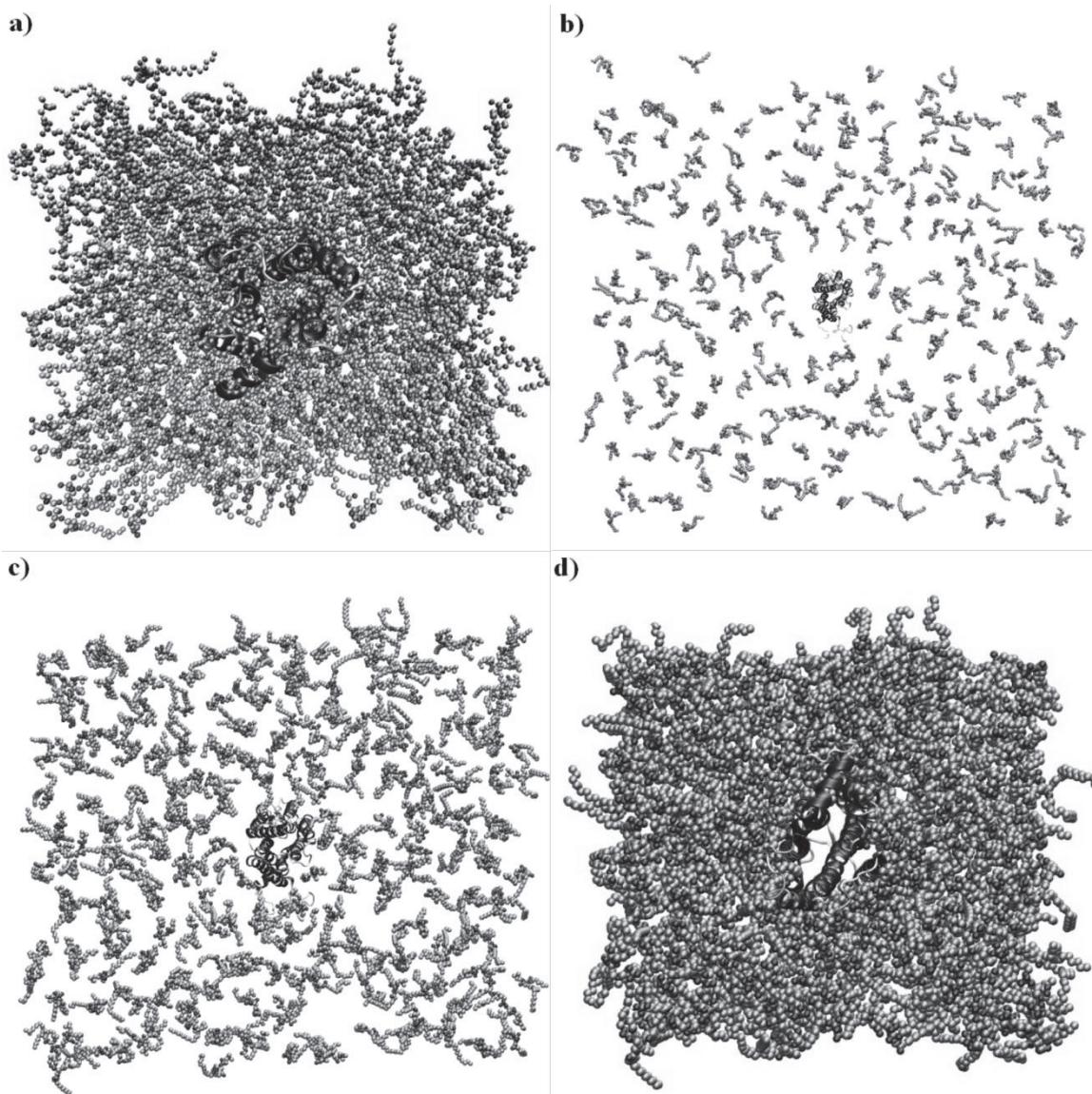


Figure 5. Inflategro method. a) Alignment of the OR on the membrane. b) Inflation of the lipid. c) Tenth step of deflation. d) Final model.

7. Solvate your system with water molecules and add ions to neutralize the system using a default protocol.
8. Maestro (Schrödinger inc.) has a membrane building protocol that allows inserting your OR in a box containing the membrane, the water phase and the ions. It is very user-friendly but requires longer equilibration time by means of molecular dynamics.

3.6. Molecular dynamic simulations

The relaxation of the system implies that its energy is minimized and then heated to physiological temperature, subsequently allowing atoms and molecules to move in their environment.

1. Minimize the energy of the system in two steps. First, minimize the energy of the solvent by freezing the OR and the lipids. Second, minimize the energy of the whole system.

3. Heat the system gradually up to the desired temperature (generally 310 K), keeping the volume of the simulation box fixed.
4. Equilibrate the system at one atm. with a semi-isotropic pressure scaling.
5. The molecular dynamics simulation production phase is continued from the last equilibration step, to collect sampling for further analysis. For most odorants, the cavity created during the homology modeling is too large and tightens during the equilibration steps. The molecular dynamics production phase is then performed to evaluate the residues in contact with the odorant from a statistical point of view. These residues can be different from those found in the starting structure obtained with the docking procedure. Fig. 6 illustrates the structure of hOR1G1 just after the docking step and after a 20 ns molecular dynamics simulation. Additionally, binding free energy estimation protocols can be considered to compare the computed affinity with experimental data, such as calcium imaging or dissociation constant measurements.

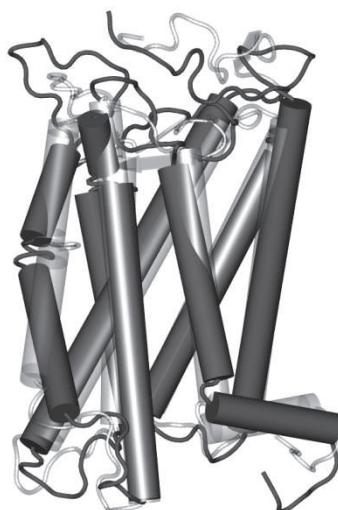


Figure 6. Difference between the initial structure (transparent), obtained after the homology building procedure and the structure relaxed after 20ns of simulation (in dark gray).

3.7. Docking rescoring

Since the interaction between an odorant and an olfactory receptor are mainly hydrophobic, one can observe a reorganization of the ligand in the binding cavity during MD simulations.

To decipher the binding mode of a ligand, multiple molecular dynamics protocol can be used. According to available computational resources, several poses obtained from the docking protocol are chosen as starting points for MD simulations. Each trajectory is then analysed using for example a MM-GBSA protocol. It has the advantage of providing a decomposition of the free energy of binding on a per-residue basis.

Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65

This method enhances the sampling of the binding cavity by the ligand, and thus allows finding the main amino acids involved in the interaction. It appears to be particularly efficient in the case of nondirectional interaction between the ligand and the cavity.

For a detailed protocol of using MM-GBSA with AMBER, the reader should refer to <http://ambermd.org/tutorials/advanced/tutorial3/>.

4. Notes

1. Even if no OR structure is known, it is important to put several other ORs in the alignment, to be sure that the crucial conserved sequences or residues are correctly accounted for during the alignment process.
2. At this stage, one can eventually perform a geometry optimization of the residues side-chains (and only the side-chains) with molecular modeling software. Indeed, it is difficult to estimate side-chain orientation.
3. If many models fulfil only a part of the criteria (discussed in Subheading 3.3, steps 3 and 4), this can point to a bad alignment of the sequences. This may necessitate a modification of the alignment (Subheading 3.1).
4. You can use a standard docking protocol for this step (not described here). Generally, the olfactory receptor is considered rigid and the only the odorants can undergo conformational changes. One can consider that the binding site will be identical to those found in rhodopsin and adrenergic receptors.
5. Membrane embedding can be done by several methods. Reference[28] summarizes them.
6. It is very useful to use a script at this stage. Indeed, the deflation of the system takes lot of simulation steps (depending on your scaling factor during the inflation).

Acknowledgments

The CINES is acknowledged for providing computer time for the project cmi1024.

References

- [1] H. Breer, Odor recognition and second messenger signaling in olfactory receptor neurons, *Semin Cell Dev Biol* 5 (1994) 25-32.
- [2] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial Receptor Codes for Odors, *Cell* 96 (1999) 713-723.
- [3] K.J. Rossiter, Structure-Odor Relationships, *Chem. Rev.* 96 (1996) 3201-3240.
- [4] D. Cozzetto, A. Tramontano, Advances and Pitfalls in Protein Structure Prediction, *Curr Protei Pept Sci* 9 (2008) 567-577.
- [5] R. Abrol, J.K. Bray, W.A. Goddard, Bihelix: Towards de novo structure prediction of an ensemble of G-protein coupled receptor conformations, *Proteins* 80 (2012) 505-518.
- [6] L. Charlier, J. Topin, C. Ronin, S.K. Kim, W.A. Goddard, 3rd, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cell Mol Life Sci* 69 (2012) 4205-4213.
- [7] T. Okada, M. Sugihara, A.-N. Bondar, M. Elstner, P. Entel, V. Buss, The Retinal Conformation and its Environment in Rhodopsin in Light of a New 2.2Å Crystal Structure, *J Mol Biol* 342 (2004) 571-583.
- [8] V.-P. Jaakola, M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y.T. Chien, J.R. Lane, A.P. IJzerman, R.C. Stevens, The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist, *Science* 322 (2008) 1211-1217.
- [9] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G.W. Leslie, C.G. Tate, G.F.X. Schertler, Structure of a β 1-adrenergic G-protein-coupled receptor, *Nature* 454 (2008) 486-491.
- [10] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.-J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka *et al*, High-Resolution Crystal Structure of an Engineered Human β 2-Adrenergic G Protein-Coupled Receptor, *Science* 318 (2007) 1258-1265.
- [11] B. Wu, E.Y.T. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi *et al*, Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, *Science* 330 (2010) 1066-1071.
- [12] E.Y.T. Chien, W. Liu, Q. Zhao, V. Katritch, G. Won Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov *et al*, Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist, *Science* 330 (2010) 1091-1095.
- [13] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han *et al*, Structure of the human histamine H1 receptor complex with doxepin, *Nature* 475 (2011) 65-70.
- [14] M.A. Hanson, C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer *et al*, Crystal Structure of a Lipid G Protein-Coupled Receptor, *Science* 335 (2012) 851-855.
- [15] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, *Genome Biol* 2 (2001) 1-12.
- [16] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Sci* 13 (2004) 240-254.
- [17] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *J Struct Biol* 159 (2007) 400-412.

PARTIE 2 : Modélisation moléculaire des récepteurs olfactifs

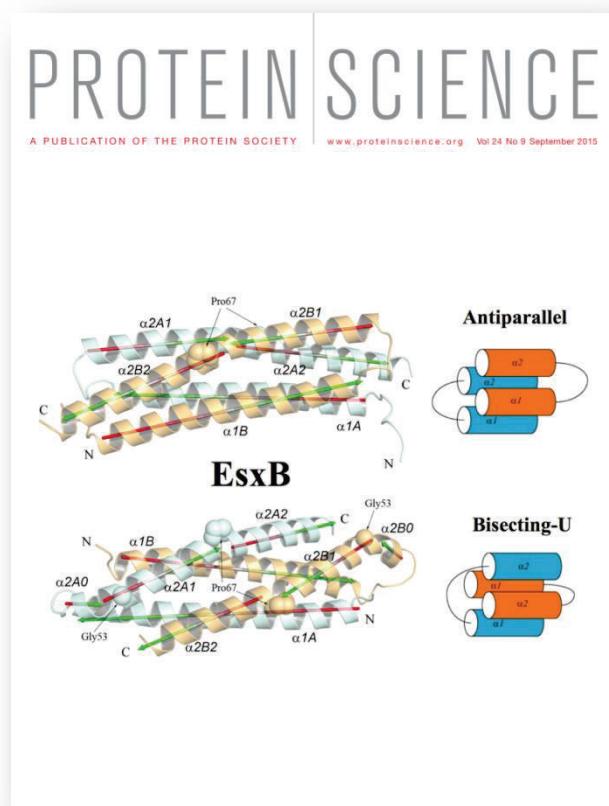
Article 3 – Charlier, Topin, de March et al. Methods in Molecular Biology 1003 (2013) 55-65

- [18] L. Doszczak, P. Kraft, H.-P. Weber, R. Bertermann, A. Triller, H. Hatt, R. Tacke, Prediction of Perception: Probing the hOR17-4 Olfactory Receptor Model with Silicon Analogues of Bourgeonal and Lilial, *Angewandte Chemie International Edition* 46 (2007) 3367-3371.
- [19] S.E. Hall, W.B. Floriano, N. Vaidehi, W.A. Goddard, Predicted 3-D Structures for Mouse I7 and Rat I7 Olfactory Receptors and Comparison of Predicted Odor Recognition Profiles with Experiment, *Chemical Senses* 29 (2004) 595-616.
- [20] T. Abaffy, A. Malhotra, C.W. Luetje, The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor: A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES, *J Biol Chem* 282 (2007) 1216-1224.
- [21] W.B. Floriano, N. Vaidehi, W.A. Goddard, M.S. Singer, G.M. Shepherd, Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor, *Proc Natl Acad Sci U S A* 97 (2000) 10712-10716.
- [22] M.S. Singer, Analysis of the Molecular Basis for Octanal Interactions in the Expressed Rat I7 Olfactory Receptor, *Chemical Senses* 25 (2000) 155-165.
- [23] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The Mouse Eugenol Odorant Receptor: Structural and Functional Plasticity of a Broadly Tuned Odorant Binding Pocket, *Biochemistry* 50 (2011) 843-853.
- [24] P. Hummel, N. Vaidehi, W.B. Floriano, S.E. Hall, W.A. Goddard, Test of the Binding Threshold Hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93, *Protein Sci* 14 (2005) 703-710.
- [25] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angewandte Chemie International Edition* 51 (2012) 1274-1278.
- [26] V.A. Avlani, K.J. Gregory, C.J. Morton, M.W. Parker, P.M. Sexton, A. Christopoulos, Critical Role for the Second Extracellular Loop in the Binding of Both Orthosteric and Allosteric G Protein-coupled Receptor Ligands, *J Biol Chem* 282 (2007) 25677-25686.
- [27] J.F. Nagle, S. Tristram-Nagle, Structure of lipid bilayers, *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* 1469 (2000) 159-195.
- [28] C. Kandt, W.L. Ash, D. Peter Tieleman, Setting up and running molecular dynamics simulations of membrane proteins, *Methods* 41 (2007) 475-488.

Article 4:

G Protein-Coupled Odorant Receptors: from sequence to structure

Claire A. de March, Soo-Kyung Kim, Serge Antonczak, William A. Goddard III & Jérôme Golebiowski,
Protein Science 24 (2015) 1543-1548



Keywords: odorant receptor, sequence, structure, binding cavity, olfactory

Abstract

Odorant receptors (ORs) are the largest sub-family within Class-A G Protein-Coupled Receptors (GPCRs). No experimental structural data of any OR is available to date and atomic-level insights are likely to be obtained by means of molecular modeling. In this article, we critically align sequences of ORs with those GPCRs for which a structure is available. Here, an alignment consistent with available site-directed mutagenesis data on various ORs is proposed. Using this alignment, the choice of the template is deemed rather minor for identifying residues that constitute the wall of the binding cavity or those involved in G-protein recognition.

Introduction

Odorant molecules are perceived by mammals through extraordinary subtle mechanisms, notably involving odorant receptors (ORs).[1] In human, the family of genes coding for ORs is one of the largest, as it represents more than 2% of our genome. At the protein level, ORs account for more than 4% of our proteome and constitute the largest sub-family of Class-A (or Rhodopsin like) G Protein-Coupled Receptors (GPCR). GPCRs are seven-transmembrane domain (7 TM) proteins that transmit extracellular signals across the plasma membrane. Although structures of some Class-A members have been experimentally solved, no experimental structure is to date available for any OR. For now, molecular modeling appears as the only way to propose atomic-level mechanisms of either ligand selectivity or receptor activation for these proteins on a structural basis. Models can either be made *ab-initio* or based on sequence homology with respect to known experimental structures.[2, 3] In both cases, sequence alignment between the candidate receptor and the experimentally determined templates is undoubtedly the crucial step.

Within the motifs that represent hallmarks of Class-A GPCR, most are shared by ORs,[4] suggesting rather similar activation mechanism upon ligand binding and similar signal transduction. It follows that templates available for now may be sufficiently adapted to recover trustable OR models. Nevertheless, ORs conserved motifs are either broader or different than those observed in Class-A GPCRs. These motifs within OR sequences are as follows, with those shared by non-olfactory Class-A GPCRs written in bold:

- **GN** in Trans-Membrane domain 1 (TM1),
- LHxPMYFFLxx**LSxxD** in TM2,
- **MAYD(E)RYV**AICxPLxY in TM3,
- **SY** in TM5,
- KAFSTCxSH in TM6,
- PxLNPxIYSLNR in TM7.

Although TM1, 2, 3 and TM7 motifs are sufficiently conserved to lead to unambiguous alignments, TM4, 5 and 6 cases are more subtle and require additional data, ideally brought by experiments. An accurate sequence alignment will provide extremely useful information on residues forming the binding cavity or involved in receptor activation. Based on a thorough alignment and analysis of conservation thresholds between mouse and human OR, such information was inferred and allowed identifying residues that contribute to ligand binding.[5] In this article, we revisit and update this data by recapitulating available experimental results published so far. We combine information gained by sequence alignments and *in vitro* data using site-directed mutagenesis to provide an optimal sequence alignment consistent with experiment. In a second step, we use this alignment to assess the choice of the template for building a representative OR and to confirm that site-directed mutagenesis data can be interpreted on a structural basis using this model.

Results

Olfactory and non-Olfactory GPCR alignment

Alignments of TM1, TM2, and TM3 sequences are straightforward as the conserved motifs in each of these TM domains are clearly identified between ORs and available GPCR structures. Figure 1 recapitulates the alignment for ORs with available site-directed mutagenesis data. In TM1, the typical Class-A GPCR ‘GN’ motif is conserved at 90 and 99% within human and mouse OR, respectively.[6, 7] Here, residue N is referenced as $N^{1.50}$, according to the Ballesteros-Weinstein notation.[8] In TM2, the PMY motif found in ORs has no equivalence in any other Class-A GPCRs but the highly conserved LSxxD in ORs is straightforward to align with the highly conserved GPCR LAxAD ($D^{2.50}$) motif. The alignment of TM3 is the easiest case because of the presence of both the D(E)RY motif ($R^{3.50}$) involved in the activation of all Class-A GPCRs, and the cysteine residue $C^{3.25}$ involved in the cysteine bridge with the Extracellular Loop 2 (ECL2). Within TM4, the tryptophan residue ($W^{4.50}$) strongly conserved in non-olfactory GPCRs is also present in ORs, with conservation of 58% and 50% within human and mouse ORs, respectively. This residue provides a good anchoring point for fitting TM4 sequences of ORs and non-olfactory GPCRs. Before considering TM5 and TM6, we focus on TM7, where the NPxxY ($P^{7.50}$) motif is conserved in all Class-A GPCRs making easy the alignment of TM7. In TM5, the highly conserved proline ($P^{5.50}$) in Class-A GPCR[8] is moderately represented in OR (conservation of 39% and 37%, in human and mouse ORs, respectively). However, the tyrosine residue of the ‘SY’ ($Y^{5.58}$) motif is strongly conserved in both GPCR sub-families (100% and 93% in mouse and human ORs, respectively). Taking this tyrosine residue as a reference assesses the accurate alignment of TM5 and remains consistent with the position of the proline residue ($P^{5.50}$) between OR and sequences associated to available X-ray structures.

TM6 is even much trickier, as this TM lacks the CWxP ($P^{6,50}$) motif considered as the TM6 hallmark of Class-A GPCR. In TM6, ORs sequences show a highly conserved KAFSTCxSH motif for which the equivalence with non-olfactory GPCR is not obvious. A ‘KA’ motif can however be identified in non-olfactory GPCRs, and a 29% conserved Proline in human ORs is aligned with the $P^{6,50}$, assessing our alignment.

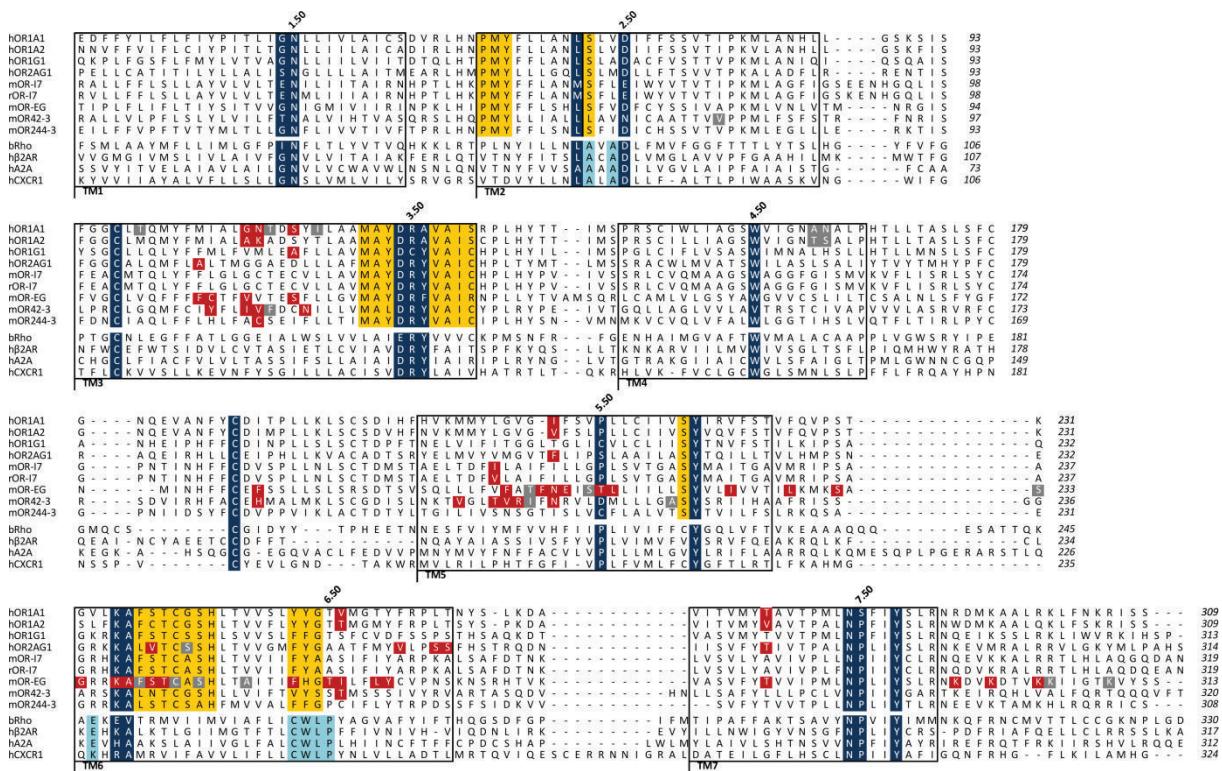


Figure 1. Alignment of ORs with some G Protein-Coupled Receptors (GPCRs). Only ORs for which site-directed mutagenesis combined to molecular modeling was available are considered. Residues commonly conserved between ORs and non-OR GPCRs (dark blue), specific to ORs only (yellow), and specific to non-OR GPCRs only (light blue) are identified. Residues which experimentally modify the OR response upon odorant stimulation are shown in red, while those which do not change the OR response are in gray. Each transmembrane (TM) domain is boxed and the Ballesteros-Weinstein numbering scheme is indicated for Class-A GPCR. An alternative numbering scheme is proposed for the TM5 and TM6 of OR, which takes into account for highly conserved residues within these TMs (orange, italicized). Site-directed mutagenesis data are reported for the Human (h) OR1A1 and hOR1A2,[9] hOR1G1,[10] hOR2AG1,[11] Rat (r) and Mouse (m) I7,[12] mOR-EG,[13, 14] mOR42-3,[15] and mOR244-3.[16] OR sequences are aligned with sequences of Bovine Rhodopsin (bRho), human β 2-adrenergic (hB2AR), human Adenosine-2A (hA2A), and human Chemokine-1 (CXCR1) receptors.

Intra and extra-cellular loops are also of importance for the function of a receptor. Here, we notably focus on ECL2 since it is involved in ligand binding and receptor structure. A disulfide bridge between ECL2 and C^{3,25} at the top of TM3 is common to all Class-A GPCRs. In ORs, three cysteines are present in ECL2 domain and one at the top of TM3, suggesting the presence of two disulfide bridges. Indeed,

in addition to the canonical S-S bridge (between C97^{3.25} and C179^{ECL2}), identification of an additional S-S bridge within ECL2 (between C169^{ECL2} and C189^{ECL2}) was characterized by mass spectrometry in hOR1D2.[17] Forcing the alignment of the canonical cysteine bridge between ORs and non-olfactory GPCRs (C97^{3.25}-C179^{ECL2}) provides a crucial data for the optimal alignment of ECL2.

This sequence alignment does not contain any gap within TM domains. The only gaps are set within loop sequences, consistent with a larger sequence and structure variability within loops with respect to the bundle.[18] Based on the alignment of Figure 1, we next address the choice of template used for building a structural model consistent with site-directed mutagenesis data.

3D structure and comparison with experimental data

Here, we analyze the accuracy of the alignment by translating it into atomic-level models. Five models of the human OR1G1 are built either with Modeller[19] using different receptor structures as templates (Bovine Rhodopsin, Human β2-adrenergic, Human Chemokine-1, and a combination of them three) or by means of the *ab initio* GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) complete sampling method.[3, 20, 21]

Figure 2 gathers information inferred from these models. Focusing on the helical TM domains, all structures are similar with Cα Root Mean Square deviations (RMSd) lower than 3 Å (see Figure 2C) between pairs of models, at the exception of that based on the chemokine receptor. The latter exhibits a RMSd value of ~ 6 Å with respect to other structures. The main difference when using the Chemokine receptor template appears for TM1, TM2 and TM7 which show a small deviation with respect to other templates. This difference has however a small influence on the position of residues lining the binding cavity. Focusing on eight of them (104^{3.32}, 108^{3.36}, 202^{5.42}, 206^{5.46}, 252^{6.48}, 256^{6.52}, 260^{6.56}, and 279^{7.42}, *vide infra*), we compute a Cα RMSd of 3.2 Å between the multi-template model and that build with the Chemokine receptor. Importantly, despite these tertiary structure weak dissimilarities, all models exhibit similar secondary folds. Furthermore, residues that constitute the wall of the binding cavity and those involved in the signaling pathway through a contact with the G-protein appear to be located in the same regions.[22, 23] As observed in all Class-A GPCRs, the canonical binding site is made up by residues belonging to TM3, TM5, TM6, and TM7.[5] Inspection of TM3 3D-structure shows that side-chains of residues 109^{3.37}, 108^{3.36}, 105^{3.33}, and 104^{3.32} participate to the binding cavity. This is consistent with a modification of the odorant response when tested in mutants expressed *in vitro* (Figure 1). In the models, residue 112^{3.40} is located under the binding cavity. Its non-synonymous mutation is consistent with a general decrease of the OR response to odorants in hOR1G1 (Ala → Ser),[24], mOR-EG (Ser → Ala or Val),[13, 14], mOR42-3 (Val → Ser),[15] and hOR1A1 (Ser → Ala).[9]

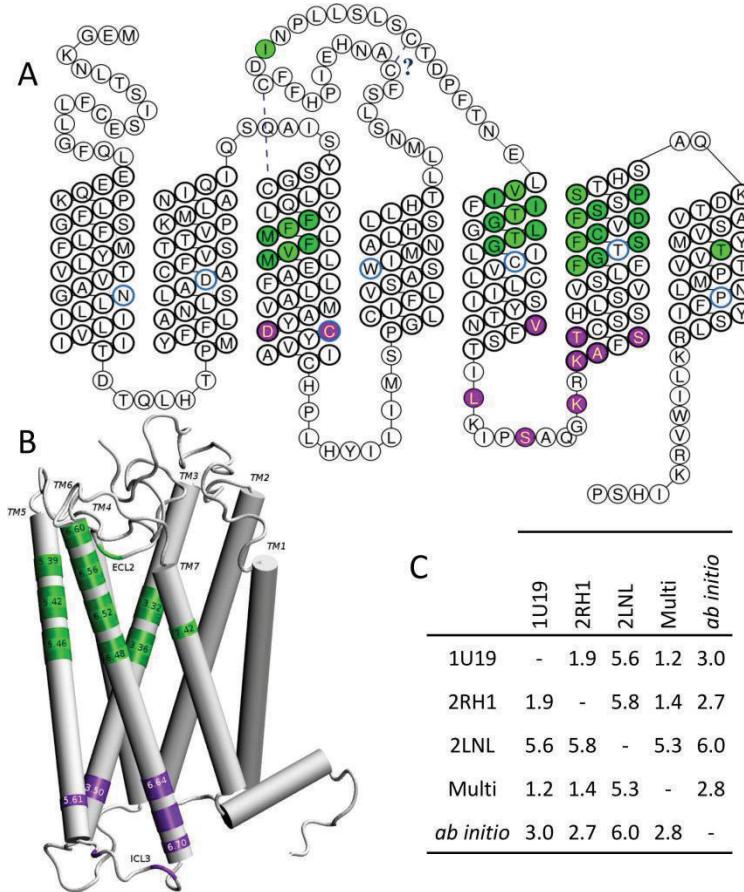


Figure 2. Residues governing the function of mammalian ORs projected onto the sequence and the structure of hOR1G1. A, snakeplot of the OR sequence with residues involved in odorant contact in green and those involved in the OR activation through a contact with the G Protein in purple. Residues in light green will be strongly in contact with the odorant, those in dark green contribute to the wall of the binding cavity. Number 50 residue of the Ballesteros-Weinstein notation are circled in blue. The cysteine bridges are also indicated. B, position of important residues on the structure of the receptor, with some Ballesteros-Weinstein notations. C, C- α positions Root Mean Square deviation (in Å) between models build using Bovine Rhodopsin (1U19), β 2-adrenergic (2RH1), Chemokine-1 (2LNL) receptor, or a multi-template (Multi) of the three receptors cited above, or an *ab initio* model.

TM4 would contribute to lining the binding cavity through one or two residues located at the top of the helix. Mutations at these positions (4.55 and 4.56) however do not affect responsiveness of the receptor,[9] suggesting that this contribution is deemed rather minor.

Amino-acids belonging to TM5 largely contribute to define the binding cavity. Side-chains of residues 199^{5.39}, 202^{5.42}, 206^{5.46} point inward the cavity, consistent with a modification of the response to odorants upon mutation on mOR-EG[13, 14] and mOR42-3 *in vitro*.[15] In mOR-EG, mutations at residues located deeper into the structure (5.50 and 5.51) also affected responsiveness of the receptor when stimulated by odorants. They would rather contribute to stabilize the receptor since they correspond to positions within the sequence showing a larger conservation (Pro at ~40% at

position 5.50, Phe/Leu at 64% at 5.51, and Ile at ~85% at 5.61) than hypervariable residues found within the cavity.[5] The main contribution of TM6 to the function of the receptor stems not only from residues within the binding cavity but also from others involved in the activation. The highly conserved aromatic residue at position 6.48 (Y/F252 is conserved at ~95%) is located at the bottom of the binding cavity. One, two, and three helix turns above, residues 255^{6.51}-256^{6.52}, 259^{6.55}-260^{6.56}, and 263^{6.59}-264^{6.60} are pointing to the cavity. These positions are in line with *in vitro* data on mOR-EG,[13, 14] mOR42-3,[15] hOR2AG1,[11] hOR1A1, and hOR1A2, where the response of the receptor upon odorant stimulation is modified by mutations at these positions.[9] Deeper into the intracellular part, the ‘KAFSTCASH’ is likely to take part in the contact with the G protein upon activation, as shown on mOR-EG,[23]. The contribution of TM7 to the binding pocket is mostly coming from residue 279^{7.42}, consistent with its impact on ligand recognition on several OR *in vitro*.[9, 11, 14]

Conclusion

We have built an alignment of mammalian Odorant Receptor sequences that recapitulates available experimental data obtained by site-directed mutagenesis. More particularly, the debatable alignment of TM5 and TM6 are now consistent with data provided by several other studies. The effect of the template in the case of homology-based approaches is deemed rather minor if one is interested in identifying residues that belong to the binding cavity or those potentially involved in the coupling of a G-protein to the OR. These data provide a robust starting point for initiating mechanistic or structural studies involving odorant receptor and their complexes with ligands.

Materials and Methods

The alignment was performed with Jalview.[25] Sequences have been firstly aligned with ClustalW prior to manual adjustments. Tools of GPCRDB have been used to obtain a snakeplot. 3D models have been built either with Modeller[19] by homology modeling using a mono- or multi-template (Bovine Rhodopsin PDB:1U19, Human β2-adrenergic PDB:2RH1 and Human Chemokine-1 PDB:2LNL) or by an *ab initio* protocol with the GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) complete sampling method[21]. Visual analysis, images, and RMSd calculations have been performed with VMD.[26]

Acknowledgments

This work was supported by grants from APEX Region PACA to J. G. (OLFACTOME) and from the Fondation Roudnitska under the aegis of Fondation de France to C.A.D.M.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] T. Yarnitzky, A. Levit, M.Y. Niv, Homology modeling of G-protein-coupled receptors with X-ray structures on the rise, *Curr Opin Drug Discov Devel* 13 (2010) 317-325.
- [3] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, in: C.J. Crasto (Eds.), Olfactory Receptors, Molecular Modelling of Odorant/Olfactory Receptor Complexes, 2013, pp. 53-65.
- [4] P. Mombaerts, Seven-Transmembrane Proteins as Odorant and Chemosensory Receptors, *Science* 286 (1999) 707-711.
- [5] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Sci* 13 (2004) 240-254.
- [6] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, *Genome Biol* 2 (2001) 1-12.
- [7] X. Zhang, S. Firestein, The olfactory receptor gene superfamily of the mouse, *Nat Neurosci* 5 (2002) 124-133.
- [8] J.A. Ballesteros, H. Weinstein, in: C.S. Stuart (Eds.), Methods in Neurosciences, [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, 1995, pp. 366-428.
- [9] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *J Struct Biol* 159 (2007) 400-412.
- [10] G. Launay, G. Sanz, E. Pajot-Augy, J.-F. Gibrat, Modeling of mammalian olfactory receptors and docking of odorants, *Biophys Rev* 4 (2012) 255-269.
- [11] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angew Chem Int Ed* 51 (2012) 1274-1278.
- [12] D. Krautwurst, K.-W. Yau, R.R. Reed, Identification of Ligands for Olfactory Receptors by Functional Expression of a Receptor Library, *Cell* 95 (1998) 917-926.
- [13] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The Mouse Eugenol Odorant Receptor: Structural and Functional Plasticity of a Broadly Tuned Odorant Binding Pocket, *Biochemistry* 50 (2011) 843-853.
- [14] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural Basis for a Broad But Selective Ligand Spectrum of a Mouse Olfactory Receptor: Mapping the Odorant-Binding Site, *J Neurosci* 25 (2005) 1806-1815.
- [15] T. Abaffy, A. Malhotra, C.W. Luetje, The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor: A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES, *J Biol Chem* 282 (2007) 1216-1224.
- [16] S. Sekharan, Mehmed Z. Ertem, H. Zhuang, E. Block, H. Matsunami, R. Zhang, Jennifer N. Wei, Y. Pan, Victor S. Batista, QM/MM Model of the Mouse Olfactory Receptor MOR244-3 Validated by Site-Directed Mutagenesis Experiments, *Biophys J* 107 (2014) L5-L8.
- [17] B.L. Cook, D. Steuerwald, L. Kaiser, J. Graveland-Bikker, M. Vanberghem, A.P. Berke, K. Herlihy, H. Pick, H. Vogel, S. Zhang, Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4, *Proc Natl Acad Sci USA* 106 (2009) 11925-11930.
- [18] M. Wheatley, D. Wootten, M.T. Conner, J. Simms, R. Kendrick, R.T. Logan, D.R. Poyner, J. Barwell, Lifting the lid on GPCRs: the role of extracellular loops, *Br J Pharmacol* 165 (2012) 1688-1703.
- [19] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper, A. Sali, in: (Eds.), Current Protocols in Bioinformatics, Comparative Protein Structure Modeling Using Modeller, 2006, pp.

- [20] S.-K. Kim, W. Goddard, III, Predicted 3D structures of olfactory receptors with details of odorant binding to OR1G1, J Comput Aided Mol Des (2014) 1-16.
- [21] J.K. Bray, R. Abrol, W.A. Goddard, B. Trzaskowski, C.E. Scott, SuperBiHelix method for predicting the pleiotropic ensemble of G-protein–coupled receptor conformations, Proc Natl Acad Sci USA 111 (2014) E72-E78.
- [22] A. Kato, K. Touhara, Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization, Cell Mol Life Sci 66 (2009) 3743-3753.
- [23] A. Kato, S. Katada, K. Touhara, Amino acids involved in conformational dynamics and G protein coupling of an odorant receptor: targeting gain-of-function mutation, J Neurochem 107 (2008) 1261-1270.
- [24] G. Launay, S. Teletchea, F. Wade, E. Pajot-Augy, J.F. Gibrat, G. Sanz, Automatic modeling of mammalian olfactory receptors and docking of odorants, Protein Eng Des Sel 25 (2012) 377-386.
- [25] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2—a multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (2009) 1189-1191.
- [26] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J Mol Graphics 14 (1996) 33-38.

Relations structure-fonction des récepteurs olfactifs

Articles 5 et 6 - Lien olfactophore-code combinatoire de ROs

Au niveau physiologique, la perception d'une odeur spécifique est très probablement due à son code combinatoire d'activation de ROs. Du point de vue du chimiste, une odeur est associée à la structure ou aux propriétés physicochimiques d'une molécule. On peut supposer que l'établissement des relations entre les structures de molécules odorantes et l'odeur qu'elles déclenchent peut se faire uniquement à partir des propriétés de celles-ci. Ce concept part du principe que des molécules possédant des points structuraux ou physico-chimiques communs posséderaient des odeurs similaires, s'affranchissant ainsi de la prise en compte des ROs. Un moyen de regrouper les propriétés d'un groupe de molécules associées à une odeur identique peut passer par la construction d'une grandeur statistique, appelée olfactophore. Dans ce modèle, les différentes propriétés physicochimiques que doit posséder une molécule pour avoir une odeur donnée sont récapitulées dans l'espace.

Cette méthode est inspirée de la recherche en pharmacologie. Les structures des composés étant connus pour agir sur une cible thérapeutique (potentiellement un RCPG) sont comparées. Le but est de déterminer les points communs de leur structure, qu'on imagine interagir avec la cible biologique. Les composés possédant ces propriétés sont suspectés d'avoir l'activité pharmacologique recherchée. De nouvelles molécules thérapeutiques peuvent être ainsi conçues de façon rationnelle, par analogie avec des molécules connues.

Dans les cas des cibles pharmacologiques, cette étape est facilitée par la ressemblance du peu de molécules capables d'activer un RCPG. En revanche, la tâche semble beaucoup plus complexe dans le cas des molécules odorantes qui peuvent, bien que possédant la même odeur, se décliner en de nombreux types de structures.

Un exemple : construction de l'olfactophore de l'odeur santalée

Le bois de santal est une matière première incontournable en parfumerie. La large utilisation de son huile essentielle dans des compositions parfumées est limitée par la difficulté de la culture de l'arbre dont elle provient. En effet, un arbre n'est exploitable pour la production qu'à l'âge de 30 ans et est très sensible aux maladies et aux insectes nuisibles. La recherche de composés possédant son odeur caractéristique est donc très active dans la communauté des chimistes.

La conception d'un olfactophore de l'odeur santalée peut guider cette recherche. Il est nécessaire de le créer à partir d'un ensemble de molécules santalées variées afin d'obtenir un modèle représentatif de tout l'espace chimique d'odorants concernés. Néanmoins, dans le cas de cette odeur, les

composés possèdent des structures chimiques extrêmement diverses (Figure 1). Comment trouver les points communs entre les structures de ces trois familles ?

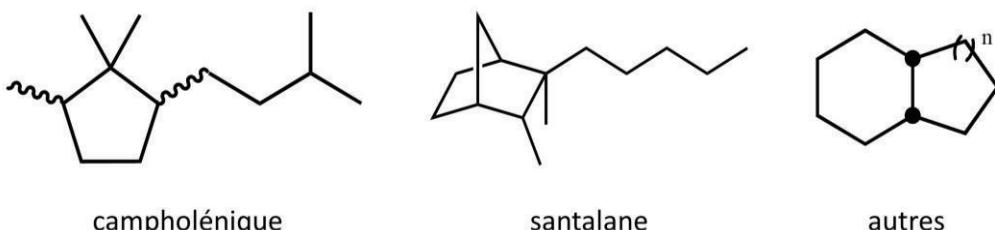


Figure 1. Structures représentatives des grandes familles de composés à odeur santalée. De gauche à droite, les structures dites campholéniques sont des cycles à cinq atomes de carbones inspirées du campholenal et possédant une chaîne latérale variable. Les molécules de type santalane sont composées d'un bicyclette 2,2,1 à 5 et 6 atomes de carbones. La nature de leur chaîne latérale est également variable. La dernière famille regroupe les composés polycycliques pouvant porter de 2 à 4 cycles consécutifs.

Dans l'article « Structure-Odor Relationships of hemisynthetic β -santalol analogues » (article 5), une bibliothèque de molécules analogues à une molécule à odeur santalée a été créée grâce à la synthèse organique. L'odeur de ces molécules a été évaluée par un panel, ce qui a permis de discriminer les molécules à odeur santalée des non santalées. La richesse de cette bibliothèque de composés, combinée aux données de la littérature, a permis de construire des modèles d'olfactophores de l'odeur santalée. Typiquement, dans ce cas où les structures chimiques sont très diverses, la stratégie employée est de créer un modèle performant pour chaque famille de composés. Trois hypothèses complémentaires d'olfactophores ont été obtenues et pourront être utiles pour la conception rationnelle d'odorants santalés. On remarquera que dans ce type d'approche, uniquement basée sur un ensemble de molécules odorantes, la physiologie de l'olfaction n'est pas prise en compte directement. On peut donc se demander dans quelle mesure les modèles obtenus ont un lien avec les protagonistes biologiques impliqués dans notre sens de l'odorat.

A quoi correspondent les olfactophores de l'odeur santalée ?

Pour le cas de l'odeur santalée, il n'est pas possible de créer un seul et unique modèle d'olfactophore performant capable de rassembler les propriétés de toutes les molécules odorantes de cette famille. En s'appuyant sur nos connaissances en physiologie de l'olfaction, on soupçonne que le code combinatoire associé à cette odeur ne peut pas être constitué d'un seul récepteur olfactif. Partant du principe que les molécules odorantes santalées interagissent avec plusieurs récepteurs, dans quelle mesure les modèles d'olfactophores conçus précédemment correspondent aux cavités de ROs impliqués ?

Articles 5-6

La perspective de l'étude est de trouver un lien entre d'une part, les modèles d'olfactophores obtenus, uniquement basés sur la structure des molécules santalées et d'autre part le pharmacophore de ROs, construit à partir de ses agonistes connus. De manière intrigante, l'un des modèles d'olfactophores se superpose quasi-parfairement avec le pharmacophore d'un RO humain, OR1G1.

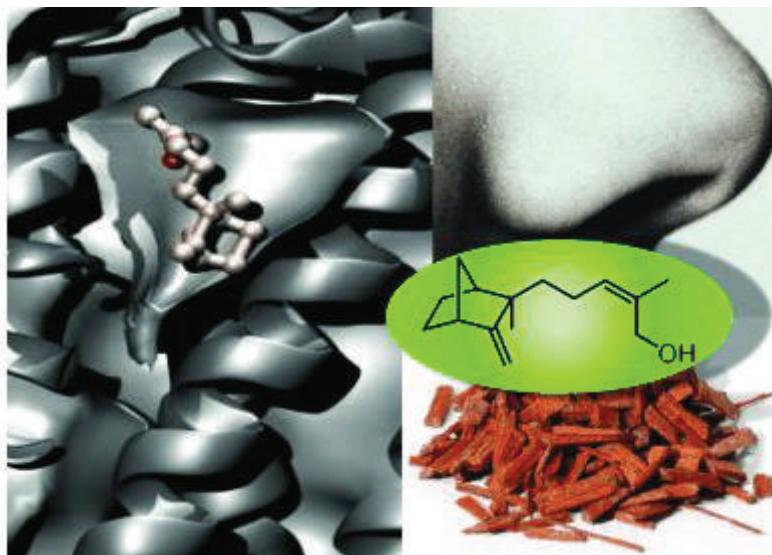
Ceci permet d'émettre l'hypothèse que les cavités des ROs impliqués dans le code combinatoire d'une odeur seraient encodées dans les modèles d'olfactophores associés à celle-ci. A travers les résultats de cette étude et les perspectives associées, nous espérons établir l'existence d'un lien entre les olfactophores, couramment conçus par les chimistes avec l'aide de parfumeurs, et les pharmacophores de ROs réalisés par les biologistes. Cette méthode pourrait assister la sélection de ROs impliqués dans le code combinatoire d'une odeur ciblée.

Ma contribution à l'article 5 « Structure-Odor Relationships of hemisynthetic β -santalol analogues » est la constitution des modèles d'olfactophores de l'odeur santalée. Dans l'article 6, j'ai réalisé l'ensemble des expériences de modélisation moléculaire et aidé aux travaux de génie génétique à l'INRA de Dijon.

Article 5:

Structure-Odor Relationships of hemisynthetic β -santalol analogues

Céline Delasalle, Claire A. de March, Uwe J. Meierhenrich, Hugues Brevard, Jérôme Golebiowski, and Nicolas Baldovini, Chemistry and Biodiversity 11 (2014) 1843-60



Keywords: Sandalwood, β -santalol, Structure-Odor Relationships, Olfactometry, Olfactophore.

Abstract

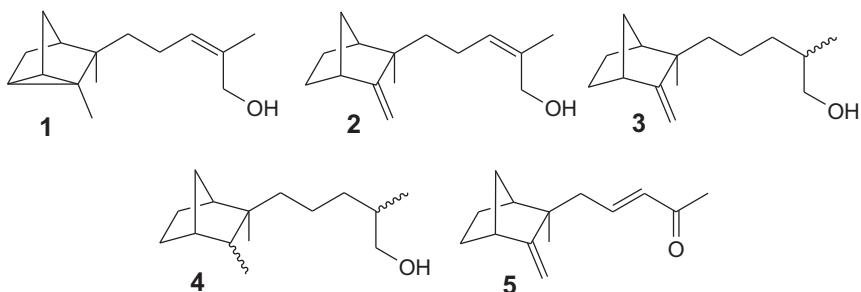
A series of 11 β -santalol analogues, including 9 new derivatives, was prepared by hemisynthesis from natural (–)-(Z)- β -santalol and studied by GC-Olfactometry to characterize their olfactory properties and potency. These compounds and 45 others selected in the literature were used to build three olfactophores by molecular modelling. Three models are obtained that gather structural and physicochemical constraints that will be useful for further design of new sandalwood odorants

Introduction

The human olfactory system is able to distinguish odorants showing very subtle differences in their chemical structure, while on the other hand, completely different molecules can present almost similar odors.[1] The ability to capture such subtleties is probably rooted in the combinatorial code that governs the perception of smell.[2, 3] Despite several attempts, the atomic-level mechanisms governing the selectivity of the sense of smell and underlying structure-odor relationships (SOR) are far from being elucidated. Therefore, new SOR studies on large sets of molecules are still of importance. They remain one of the best tools in the quest to unravel the functioning of the olfactory system. SOR of sandalwood odorants[4-6] are highly interesting and challenging since many structurally unrelated compounds present the very typical and characteristic woody fragrance of natural sandalwood, which is one of the most precious natural ingredient used in perfumery.

East Indian Sandalwood essential oil (EO) is produced by steam distillation of the heartwood and roots of *Santalum album*, an evergreen hemiparasitic tree native from southern India.[7, 8] The cultivation of *S. album* is intended almost exclusively for the perfume industry, and is complicated by the slow growth of the trees which must be at least 30 years old to be exploitable for EO production. Moreover, the trees are sensitive to pests and diseases and are frequently victims of poaching.[9] For all these reasons, the price of East Indian *Santalum album* EO has grown continuously over the years, up to an average price around 1500 €/kg in 2012. Several other *Santalum* species are now cultivated for EO production, like New Caledonian *S. austrocaledonicum*[10] or Australian *S. spicatum*;^[11, 12] they are interesting substitutes, but *Santalum album* still remains the most esteemed natural sandalwood.[9] The major components of the EO are sesquiterpenic alcohols, mainly (+)-(Z)- α -santalol 1 and (–)-(Z)- β -santalol 2. The latter constitutes about a fifth of the EO, and shows the most characteristic odor, with powerful sandalwood, milky and urinous tonalities of the EO.[13] Its structure has been elucidated by Ruzicka et al. in 1935 and its absolute configuration determined by total synthesis fifty years later.[14-16] Many syntheses of racemic and enantiopure 2 have been published[17-33] but no economical process for a large-scale industrial production of synthetic β -santalol has yet been available, even if the recently reported advances in this field suggest that it may change soon.[34] Fortunately, fragrance industry made available several synthetic substitutes of

sandalwood oil thanks to the often serendipitous discoveries. Various molecules structurally unrelated to 2 were found to possess a sandalwood odor.



The first SOR on santalanes was published by Fanta et al. who prepared diastereomeric mixtures of dihydro- and tetrahydro- β -santalols (3, 4) in their quest for β -santalol analogues, easier to synthesise than 2.[35, 36] The odor of 4 was devoid of any sandalwood character, but 3 “retained the strong and characteristic sandalwood note”. From then on, numerous analogues of sandalwood odorants have been synthesised and their olfactory profiles reported. Thus acquired SOR data were used in several attempts - often with help of molecular modelling - to generate sandalwood olfactophore models.[5, 6, 37-41] These models usually contain three general features of the osmophoric pattern: a hydroxyl group, a lipophilic group (~3 Å distant from the hydroxyl group) and a bulky rigid group (4-7 Å distant from the hydroxyl group).[39, 41, 42] In spite of the large number of β -santalol analogues included in these works, even the latest models are still unable to explain the olfactory properties of all sandalwood odorants and their structural analogues. Additional data should then be accumulated without restricting them to alcohols as in most of published studies. As a matter of fact, several sandalwood odorants bear other functionalities (aldehydes, ketones, nitriles)[43-49] and their consideration has been recommended in order to gain a deeper understanding of structural requirements for sandalwood odor.[6]

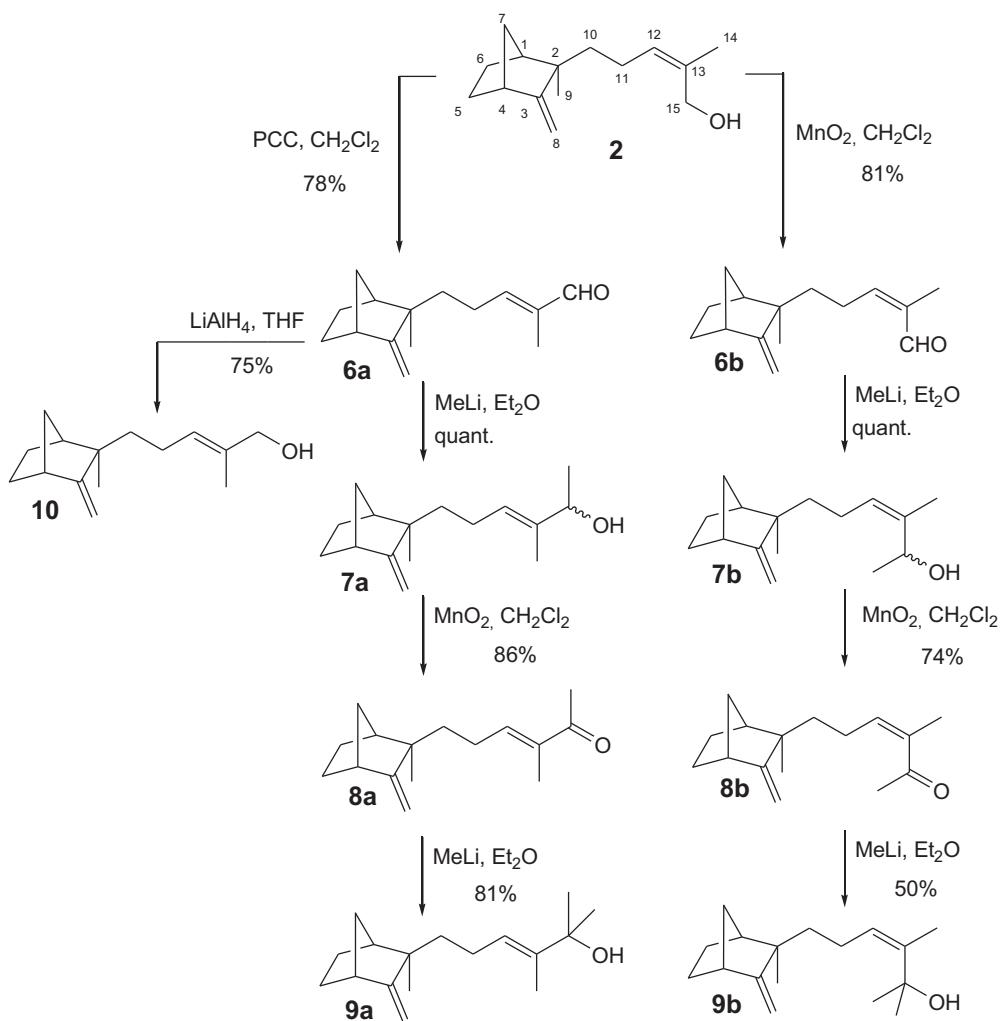
A major issue encountered in SOR investigations on santalane derivatives is due to the flexibility of the side chain bearing the alcohol function. The large number of low energy conformers hampers identifying optimal molecular geometries. In such cases, the comparison of large series of analogues still remains the only way to establish robust SOR rules.[5] Hence, to assess the influence of this part of the molecule, we chose to collect further information by focusing on analogues of 2 differing only by the structure of the side chain. Indeed, reports on such modifications[16, 39, 41, 43, 50] are rare compared to the extremely large number of sandalwood odorants described in SOR studies.

Therefore, we performed the hemisynthesis of 12 analogues starting from natural ($-$)- β -santalol. All of these compounds were subsequently analysed by Gas Chromatography-Olfactometry (GC-O) to characterize their olfactory properties and measure their relative potency. These results and

literature data have been compiled (57 structures) and used for generation of a sandalwood olfactophore model.

Results and discussion

β -santalol analogues 6-13 were prepared by hemisynthesis from natural (–)-(Z)- β -santalol 2 isolated from East Indian *Santalum album* essential oil. The oxidation of 2 with pyridinium chlorochromate (PCC) furnished (E)- β -santalal (6a) (Scheme 1).



Scheme 1. Synthesis of β -santalol analogues.

Indeed, the double bond isomerisation during PCC oxidation has already been noticed in the original publication describing the properties of this reagent.[51] Even if both (Z)- and (E)-santalals are found in the EO,[52, 53] the only olfactory properties reported have been those of 6a, described as “sweaty, urine, sexy, sandalwood”.[43] Then, we synthesised its (Z)- isomer (6b) using manganese dioxide as an oxidant. (Z)- and (E)- α -methyl- β -santalol (7a and 7b) were prepared by addition of methylolithium on these aldehydes. Both 7a and 7b were obtained as an inseparable mixture of diastereoisomers, not resolved in our GC conditions, and showing the same ^1H and ^{13}C spectral data

with only very low chemical shift differences on some ^{13}C signals of 7b. The diastereoisomeric mixture 7a has already been synthesised by Buchbauer et al. but was wrongly described as the mixture 7b - and integrated as such in their calculations - since the (Z) \rightarrow (E) isomerization due to the reaction with PCC was not taken into account.[50] We also prepared the tertiary alcohols α,α -dimethyl- β -santalols (9a and 9b) in order to determine the influence of the steric hindrance around the hydroxyl group on the olfactory properties. 9a and 9b were obtained by MnO₂ oxidation of 7a and 7b, respectively, followed by treatment with methylolithium. The ketonic intermediates 8a and 8b were also included in the study, for comparison with nor- β -santalenone (5), a minor constituent of sandalwood EO[10, 43] possessing a “sweaty, woody, green sandalwood” fragrance.[43]

(E)- β -santalol (10) is a natural constituent of sandalwood EO, usually present in much lower concentration than 2.[10, 54-56] In the literature, the description of its olfactory character is somewhat controversial: it was reported to be less potent than 2, but either qualitatively similar[18] or woody-medicinal.[57] Therefore, to fulfil the set of primary, secondary and tertiary alcohols related to 2, we also synthesised 10 by reduction of 6a (Scheme 1).

The β -santalol analogues 6-10 were then characterized by Gas Chromatography - Olfactometry (GC-O). This technique offers two advantages: since even very low amounts of odorant impurities can significantly alter the olfactory properties of a given compound,[58] the GC separation ensures a highly efficient purification which eliminates most of the potential odorant contaminations. Moreover, by injecting successive dilutions of a given compound, the GC-O permits measurement of its olfactory potency by taking into account the concentration of the most diluted solution still perceived by half of the panellists, as usually practised in the Aroma Extract Dilution Analysis (AEDA) method.[59] In our work, 8 panelists participated to these evaluations, after checking the uniformity of their response to the odor of β -santalol 2. Such a “calibration” of the panel is crucial to ensure validity of the characterisations. Indeed, specific anosmias are frequently reported, especially with sandalwood odorants, and moreover, the qualitative description is generally strongly dependent on the olfactory culture of the evaluator. Therefore, it varies a lot as a function of the semantic descriptors of the evaluator, complicating the comparison of the data reported from different literature sources.[5] The results of the evaluation of compounds 6-13 are reported in Table 1. Primary and secondary alcohols (7a, 7b, 10) exhibit the same characteristic typical sandalwood scent of 2, and the double bond configuration affects only the potency of this note: (Z)- and (E)- β -santalols (2 and 10) possess the same intensity but 7a is four times less potent than 2, while 7b is twice more powerful. Interestingly, (Z)- β -santalal (6b) also possesses this sandalwood note, with additional woodier aspects, being albeit much less potent than 2. On the contrary, the odor of (E)- β -santalal (6a) is totally different from that of 2, but shows a peculiar sweaty tonality, already noticed by

Mookherjee et al..[43] The ketones 8a and 8b also show related sweaty characters, with the (E)-ketone (8a) being the most potent compound of the series. Surprisingly, alcohols 9a and 9b are odorless, despite the fact that many reported sandalwood odorants are tertiary alcohols.[59]

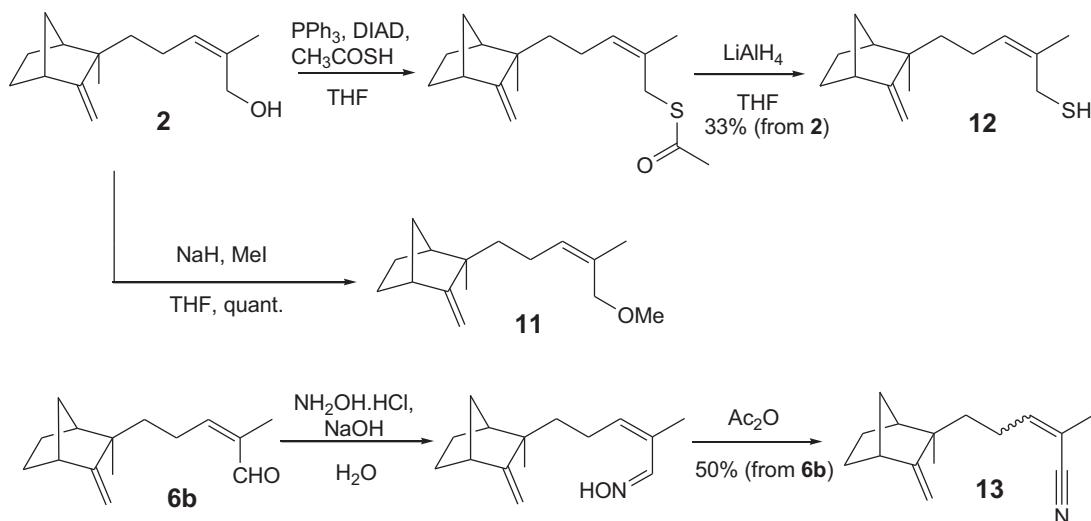
Table 1. Olfactory properties of santalane derivatives 6-11.

Compound	Qualitative Olfactory Properties	Relative Olfactory Potency ^{a)}
2	Typically sandalwood	1 (reference)
6a	Sweaty	0.125
6b	Similar to 2 , but slightly woodier	0.125
7a	Similar to 2	0.25
7b	Similar to 2	2
8a	Sweaty	8
8b	Sweaty	1
9a	Odorless	-
9b	Odorless	-
10	Similar to 2	1
11	Almost odorless	-

^{a)}) GC-O measurements, with β -santalol (**2**) taken as a reference, a value > 1 means that the

Chastrette et al. have reported the woody odor of β -santalyl methyl ether (11).[41] In Structure-Activity studies, the comparison of an alcohol with its ether is interesting since it provides insights into the role of the alcohol function, which can act either as a hydrogen bond donor or as an acceptor. Moreover, dipole moments of alcohols and ethers are slightly different. Indeed, for some fragrant sesquiterpenic alcohols of cedarwood and vetiver the conversion alcohol \rightarrow ether leads to an improvement of the olfactory quality and intensity.[60] We synthesised 11 by conventional methods. In agreement with Chastrette observation, it proved to be actually almost odorless, very slightly woody, but without any sandalwood character (Scheme 2).

Replacement of the oxygen atom by another heteroatom can provide interesting information on the receptor/odorant interactions.[1] We thus also evaluated (–)-(Z)- β -santalthiol (12). The thiol function is known to be a weak hydrogen bond donor rather than an acceptor. In addition, the thiols are well known in perfumery for their characteristic sulphury fragrance when concentrated, but often presenting fruity or floral notes when diluted.[61] 12 was synthesized in two steps from 2 by reduction of its corresponding thioacetate obtained by a Mitsunobu reaction. It shows a typical thiol note which becomes slightly woody in large dilutions, but devoid of any sandalwood aspect.

**Scheme 2.** Functional group modifications on β -santalol.

In several cases, the olfactory properties of nitriles are rather similar to those of the corresponding aldehydes[1] as with citral and Citralva®. Since we observed that the odor of (*Z*)- β -santalal (6b) was typically sandalwood, we checked if this character was retained in nitrile 13 which was prepared in two steps via the oxime obtained from 6b. This transformation led to a ca. 1:1 mixture of (*Z*) and (*E*) isomers which could be resolved and separately evaluated by GC-O. Both show only the same floral - citrus note.

Focusing on sandalwood odor perception, several models have been provided in the literature, on the basis of structural or electronic properties of the odorants.[5, 6, 37-42] For the generation of our sandalwood olfactophores, we considered a large data set of 57 sandalwood odorants and their derivatives including 11 new compounds described in this work and 46 components reported in the literature. These latter molecules were carefully selected to exclude any structure containing stereogenic centers with uncertain relative and/or absolute configurations. For these same reasons, neither 7a nor 7b, which are nearly racemic mixtures, were integrated into the santalane dataset. Among these 57 structures, 29 share the typical sandalwood odor and were considered to build the olfactophore hypothesis. The inactive compounds were then used to add exclusion volumes that correspond to areas that should contain no atoms in a sandalwood odorant. So-called ‘common feature olfactophores’ were generated with a minimum of four features, chosen between hydrogen bond donors/acceptors or hydrophobic sites. The structural diversity in our selection of 57 molecules makes the olfactophore approach difficult. It is indeed counter-intuitive to encompass the whole combinatorial code of OR activation in a single model that would be built with too many constraints. The complexity of the combinatorial code of ORs activation by sandalwood odorants was demonstrated on rat ORs. For 5 sandalwood odorants, the measured activation pattern amongst different olfactory neurons was very different.[62] How such discrimination can be achieved remains

unclear although it is obvious that the structural diversity of these odorants may lead to activation of different receptors. Their interactions with human ORs have not yet been described, and their common sandalwood character is still not attributed to any specific receptors. The combinatorial code of the mechanism of olfaction implies indeed that SOR are much more complex than their pharmacological counterparts, where a single biological receptor generally controls the activity.

Table 2. Olfactophore results obtained with santalane derivatives.

Compound ^{a)}	Fit value ^{b)}	Pharmaprint ^{c)}	Activity ^{d)}	ΔE ^{e)}	Ref
14	0.050	'111'	2	5.4	[63]
15	0.010	'111'	2	9.7	[64]
16	0.010	'111'	2	9.4	[63]
10	0.002	'111'	2	6.0	
2	0.001	'111'	2	5.1	
13b	0	'111'	0	3.0	
17	<i>no fit</i>	-	0	-	[65]
18	<i>no fit</i>	-	0	-	[66]
19	<i>no fit</i>	-	0	-	[67]
20	<i>no fit</i>	-	0	-	[67]
11	<i>no fit</i>	-	0	-	
12	<i>no fit</i>	-	0	-	
13a	<i>no fit</i>	-	0	-	
6a	<i>no fit</i>	-	0	-	
6b	<i>no fit</i>	-	2	-	
8a	<i>no fit</i>	-	0	-	
8b	<i>no fit</i>	-	0	-	
9a	<i>no fit</i>	-	0	-	
9b	<i>no fit</i>	-	0	-	
1	<i>no fit</i>	-	0	-	[56] ^{f)}
21	<i>no fit</i>	-	0	-	

a) In italics are reported molecules that were excluded by the HipHop and HipHopRefine algorithms of CATALYST during hypothesis generation. Their fit value has been obtained afterwards by fitting their structure into this model. b) The fit values range between 0 (mainly due to atoms within exclusion zones) and 3 (that would reveal a perfect fit at the center of the three features spheres, without any atom within exclusion zones). 'no fit' means that the protocol is unable to fit any conformer within the model. c) The 3 values are defined as: 1 – if a portion of a molecule matches an olfactophore feature, 0 if not. The last digit corresponds to the fit within the hydrogen bond donor sphere. d) An activity of 2 means that the odorant shows a sandalwood note, if not, the activity is 0. e) ΔE indicates the difference of energy (in kcal.mol⁻¹) between the conformer showing the best fit and the more stable conformer. f) The literature is somewhat controversial concerning the qualitative descriptors associated with the scent of 1, with many sources describing it as "sandalwood". Obviously, since AEDA studies on East Indian Sandalwood essential oil showed that 1 was the second most important contributor to the odor of the whole

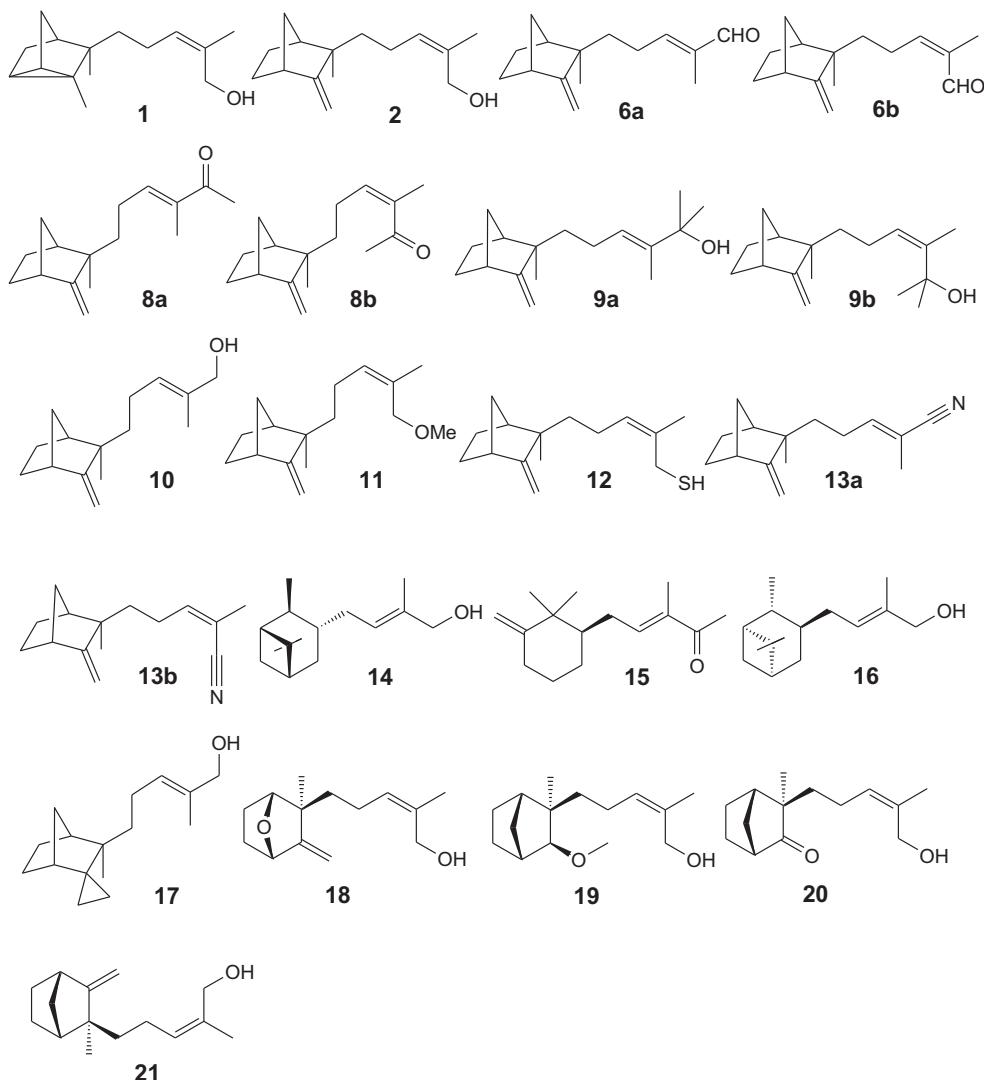
EO [44], this constituent is clearly reminiscent of some of the tonalities of natural sandalwood. However, if we define the “typically sandalwood” character as those shared by 2 and related sandalwood synthetic substitutes, then α -santalol is devoid of this typical sandalwood scent. It is indeed considered as such in our study, on the basis of an evaluation of a highly purified (>99.8%) sample by trained perfumers who reported that it possess a “relatively weak odor, slightly woody, reminiscent of cedarwood” [57].

Consequently, our set was separated into 3 structural groups, viz. santalane-like structures, campholenic derivatives and other, miscellaneous compounds. The three groups contain 21, 25 and 11 molecules, respectively (tables 3 - 5). For each group, a series of several olfactophore hypotheses were generated using CATALYST software CATALYST (Catalyst version 4.9.1 software; Accelrys Inc., San Diego, CA, August 2004). See ref. [37] for an example of utilization of this software in the field of SOR studies on sandalwood odorants. Amongst the best hypotheses of each subset, we chose the one that most accurately splits odorants between sandalwood and non-sandalwood. All three selected hypotheses are made up of 2 hydrophobic features and a hydrogen-bond acceptor as shown in fig.1. The three olfactophores differ more particularly in the position of a hydrophobic site. The large variability among these olfactophores suggests that the sandalwood perception is subjected to the activation of several olfactory receptors bearing quite different binding sites. Many exclusion volumes inferred from the inactive set of molecules restrain the space allowed for a molecule to be considered as active. 74, 50 and 17 exclusion zones are present for the santalane, campholenic and miscellaneous derivatives olfactophores, respectively.

These models are knowledge-based and have not been tested against molecules other than structures of the training sets used in their generation. They just provide statistic information on the chemical features and the distances that characterize and individualise active compounds of the database. The santalane model is closely related to that reported by Bajgrowicz and Frater,[37] with inter-feature distances reported in table 2. Other hypotheses are more different with inter-features varying between 3 and 8 Å.

Tables 3, 4 and 5 report the quantitative evaluation of the models. An analysis of raw data leads to errors of 4.8%, 24% and 9% for santalane, campholenic and miscellaneous groups, respectively. False positive candidates can nonetheless be identified through the analysis of the pharmaprint. For example, 47 in the miscellaneous group (table 5) possesses a high fit value due to an excellent fit within the hydrophilic and one of the hydrophobic features but it is totally out of the third sphere. This suggests that fulfilling the three features is mandatory to model active compounds. According to this rule, and in spite of their good fit values, 47 and 32-39 can be rationally recovered as true negatives within miscellaneous and campholenic groups, respectively. Taking into account this

analysis, the errors fall to 4.8%, 16% and 0% for santalane, campholenic and miscellaneous groups, respectively.



The presence of many exclusion zones is most of the time the cornerstone of their discriminating power. The poor fit value of non-sandalwood odorants is mainly due to the presence of atoms within these forbidden regions. 9b is a typical example: the addition of two methyl groups at the alcohol function with respect to 2 abolishes the predicted activity. These additional groups both occupy exclusion zones and hinder the fit of the molecule into the model (table 3).

Table 3. Olfactophore results obtained with campholenic derivatives

Compound	Fit value	Pharmaprint	Activity	ΔE	Ref
22	2.28	'111'	2	0.0	[68]
23	2.17	'111'	2	4.1	[68]
24	1.93	'111'	2	5.0	[69]

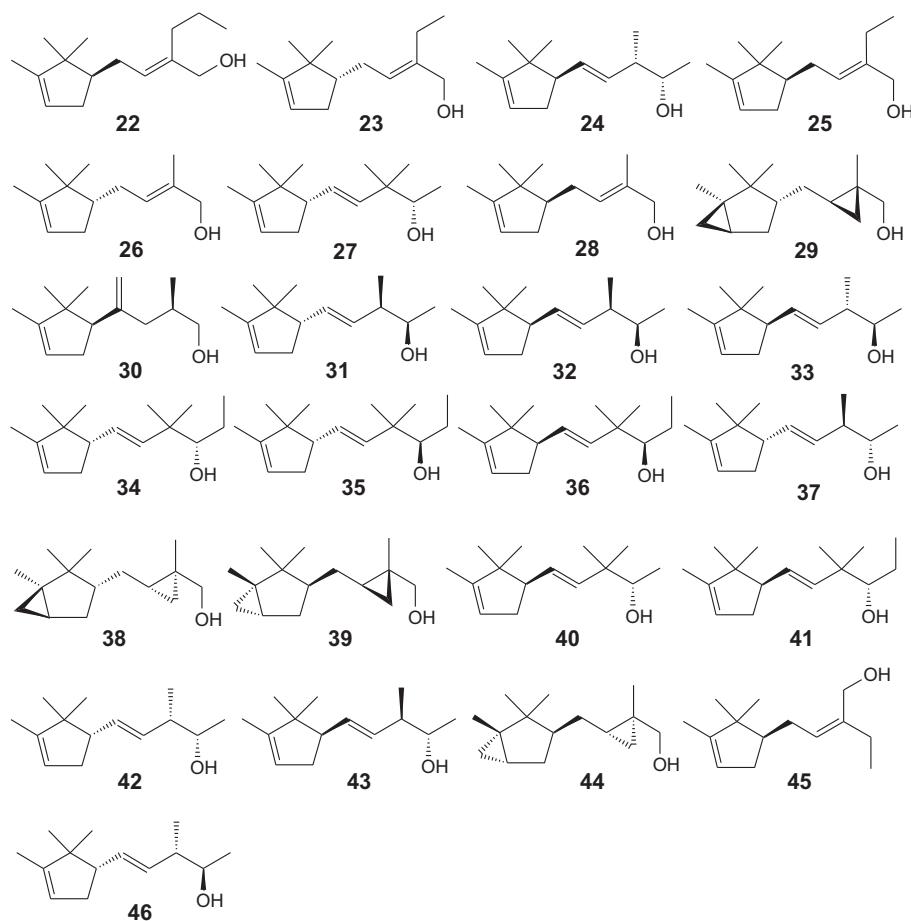
PARTIE 3 : Relations structure-fonction des récepteurs olfactifs

Article 5 – Delasalle, de March et al. 11 (2014) 1843-1860

25	1.90	'111'	2	4.3	[68]
26	1.60	'111'	2	0.0	[68]
27	1.58	'111'	2	3.8	[70]
28	1.37	'111'	2	3.6	[68]
29	1.36	'111'	0	3.7	[71], [72]
30	1.11	'111'	2	0.0	[48]
31	1.01	'111'	0	3.6	[69]
32	0.99	'101'	0	2.6	[69]
33	0.78	'111'	0	3.1	[69]
34	0.72	'111'	2	4.8	[73]
35	0.65	'111'	2	3.1	[73]
36	0.65	'111'	2	3.1	[73]
37	0.56	'111'	2	4.6	[69]
38	0.49	'111'	2	1.6	[71], [72]
39	0.48	'011'	0	4.8	[71], [72]
40	0.44	'111'	2	3.1	[70]
41	0.39	'111'	2	2.9	[73]
42	0.32	'111'	0	3.0	[69]
43	0.22	'111'	2	0.0	[69]
44	0.05	'111'	2	4.6	[71], [72]
45	0.02	'111'	0	3.0	[68]
46	0.01	'111'	0	4.7	[69]

See Table 2 for features description.

Although several chemical groups can fulfil the hydrophobic feature of each olfactophore, the presence of an oxygen atom seems mandatory to act as the hydrogen-bond acceptor or donor feature. Indeed, alternative models showed similar hydrophobic features associated with a hydrophilic donor feature. Any variation at this position leads to a non-sandalwood odor. For example, 12, a derivative of 2 where the oxygen atom is replaced by a sulphur atom is totally devoid of any sandalwood character. This suggests that an oxygen atom in a hydrogen-bond acceptor or donor group triggers the activation of most sandalwood ORs.

**Table 4.** Olfactophore results obtained with miscellaneous derivatives.

Compound	Fit value	Pharmaprint	Activity	ΔE	Ref
47	1.42	'011'	0	2.0	[74]
48	0.55	'111'	2	2.2	[74]
49	0.50	'111'	2	3.4	[74]
50	0.30	'111'	2	0.6	[74]
51	0.26	'111'	2	0.0	[75], [76]
52	0.24	'111'	2	0.0	[77]
53	0.11	'111'	2	0.5	[74]
54	0.11	'111'	0	0.0	[74]
55	0.03	'011'	0	3.0	[77]
56	0.00	'111'	0	0.0	[56], [77]
57	no fit	-	0	-	[76], [75]

See Table 2 for features description.

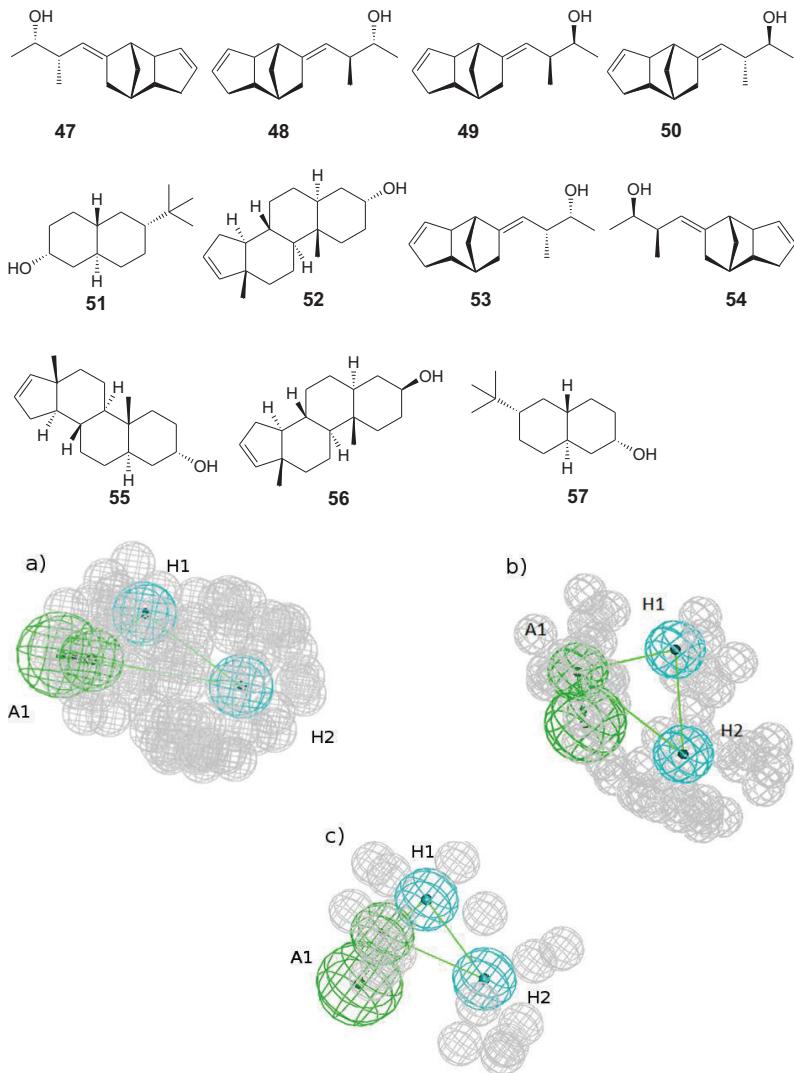


Figure. The three olfactophores inferred from the santalane (a), the campholenic (b) and the miscellaneous (c) groups. Hydrophobic features (H1 and H2) are shown in light blue, H-bond acceptors (A1) in green and exclusion volumes in grey.

Conclusions

A series of sandalwood derivative has been synthesised and evaluated in order to acquire more sandalwood structure odor relationship data. Variation among the derivatives consisted in modifications of the ($-$)- β -santalol side chain. The CATALYST common feature olfactophore approach was applied to a set of 57 compounds including our new and literature odorants. The molecules were split into three categories according to their structure (santalane, campholenic and miscellaneous). Olfactophore models for the three structural groups were built. They provide structural and physicochemical constrains potentially useful for further synthesis of sandalwood analogues. The santalane and miscellaneous models accurately rank most of the training set compounds according to their sandalwood note. Concerning the campholenic group, a detailed analysis of the pharmaprints allows a decrease of the error from 25% to 16%. The lack of sandalwood note in the

scent of thiol and nitrile derivatives of β -santalol suggests that the presence of an oxygen atom as a hydrogen bond acceptor or donor is necessary to elicit a sandalwood redolence. In the santalane group, (Z)- β -santalal (6b) is the only false negative (table 3). It should be noted that a chemical equilibrium might exist inside the cavity of the receptors, which can favor formation of derivatives. Accordingly, the hydrated form of 6b was built and tested against the santalane olfactophore hypothesis but did not fit. However, two out of its four enol forms fitted with reasonably good fit values. This observation should not be necessarily taken as an additional proof of validity of the olfactophore model, but underlines the necessity to consider the whole range of events that could convert an intrinsically odorless structure to an active derivative. Specific environment at the binding site of an OR might induce such transformations, and careful analysis of all possible odorant-OR interactions seems indispensable for a better understanding of SOR data.

Finally, the identification of multiple olfactophore models is in agreement with the postulated combinatorial coding of the sandalwood odor. Cavities of multiple sandalwood ORs are probably quite different from each other and might correspond to the spatial arrangement of features proposed by each reported hypothesis.

References

- [1] G. Ohloff, W. Pickenhagen, P. Kraft **Scent and chemistry. The molecular world of odors.** Weinheim: Wiley-VCH; 2012.
- [2] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [3] U.J. Meierhenrich, J. Golebiowski, X. Fernandez, D. Cabrol-Bass, The molecular basis of olfactory chemoreception, *Angewandte Chemie International Edition* 43 (2004) 6410-6412.
- [4] P. Kraft, J.A. Bajgrowicz, C. Denis, G. Frater, Odds and trends: recent developments in the chemistry of odorants, *Angew. Chem. Int. Ed. Engl.* 39 (2000) 2980-3010.
- [5] C. Chapuis, In *The Quest for a Virtual Pseudo Receptor for Sandalwood-Like Odorants*, *Chemistry & Biodiversity* 1 (2004) 980-1021.
- [6] K.J. Rossiter, Structure-Odor Relationships, *Chem. Rev.* 96 (1996) 3201-3240.
- [7] E.J. Brunke, W. Rojahn, *Sandalwood oil*, *Dragoco Rep.* 27 (1980) 127-135.
- [8] E. Guenther **The Essential Oils**, vol. 5, E. Guenther edn. New York: Van Nostrand Co.; 1948.
- [9] M. Nageswara-Rao, Indian Sandalwood Crisis, *Perfum. Flavor.* 33 (2008) 38-43.
- [10] N.A. Braun, M. Meier, F.-J. Hammerschmidt, New Caledonian sandalwood oil - A substitute for East Indian sandalwood oil?, *J. Essent. Oil Res.* 17 (2005) 477-480.
- [11] C. Valder, M. Neugebauer, M. Meier, B. Kohlenberg, F.-J. Hammerschmidt, N.A. Braun, Western Australian sandalwood oil-new constituents of *Santalum spicatum* (R. Br.) A. DC. (Santalaceae), *J. Essent. Oil Res.* 15 (2003) 178-186.
- [12] P. Biggs, Sustainable Australian Sandalwood, *Perfum. Flavor.* 32 (2007) 28-29.
- [13] J. Verghese, T.P. Sunny, K.V. Balakrishnan, (+)-a-Santalol and (-)-b-santalol(Z) concentration, a new quality determinant of east Indian sandalwood oil, *Flavour Fragr. J.* 5 (1990) 223-226.
- [14] L. Ruzicka, G. Thomann, Polyterpenes and polyterpenoids. XCIII. Constitution of β -santalol and β -santalene, *Helv. Chim. Acta* 18 (1935) 355-362.
- [15] H.C. Kretschmar, W.F. Erman, Total synthesis and geometric configuration of dl- β -santalol, *Tetrahedron Lett.* (1970) 41-44.
- [16] E.J. Brunke, A. Boehme, H. Struwe, Cyclopropane ring cleavage in the α -santal series. I. Absolute configurations of (-)- β -santalol, (+)-*epi*- β -santalol, and (*E*)-(-)- β -santalol, *Liebigs Ann. Chem.* (1982) 1105-1110.
- [17] G.-f. Zhong, M. Schlosser, A short and simple access to both enantiomers of epi-beta - santalene and (Z)-epi-beta -santalol, *Synlett.* (1994) 173-174.
- [18] A. Krotz, G. Helmchen, Total syntheses of sandalwood fragrances: (Z)- and (*E*)- β -santalol and their enantiomers, *ent*- β -santalene, *Tetrahedron: Asymmetry* 1 (1990) 537-540.
- [19] A. Krotz, G. Helmchen, Total syntheses, optical rotations and fragrance properties of sandalwood constituents: (-)-(Z)- and (-)-(E)- β -santalol and their enantiomers, *ent*- β -santalene, *Liebigs Ann. Chem.* (1994) 601-609.
- [20] R.L. Snowden, P. Sonnay, G. Ohloff, Stereoselective syntheses of (\pm)-*epi*- β -santalene and (\pm)-*epi*- β -santalol, *Helv. Chim. Acta* 64 (1981) 25-32.
- [21] T. Gibson, Z.J. Barneis, Approach to the synthesis of β -santalene sesquiterpenes. Isolation of exo-norbicycloekasantalal from East Indian sandalwood oil, *Tetrahedron Lett.* (1972) 2207-2210.
- [22] E.J. Corey, R. Hartmann, P.A. Vatakencherry, Synthesis of dl-santalene and dl-*epi*- β -santalene by stereo-specific routes, *J. Am. Chem. Soc.* 84 (1962) 2611-2614.
- [23] G.L. Hodgson, D.F. MacSweeney, T. Money, Synthesis of (\pm)-campherenone, (\pm)-epicampherenone, (\pm)- β -santalene, (\pm)-*epi*- β -santalene, (\pm)- α -santalene, (\pm)-ylangocamphor, (\pm)-copacamphor, and (\pm)-sativene, *J. Chem. Soc., Perkin Trans. 1* (1973) 2113-2130.
- [24] G.L. Hodgson, D.F. MacSweeney, R.W. Mills, T. Money, Synthesis and absolute configuration of the terpenes (-)-campherenone, (+)-epicampherenone, (-)- β -santalene, and (+)-*epi*- β -santalene, *J. Chem. Soc., Chem. Commun.* (1973) 235-236.

- [25] M. Bertrand, H. Monti, K.C. Huong, Stereoselective synthesis of (\pm)- β -santalene and of new unnatural santaloids, *Tetrahedron Lett.* (1979) 15-18.
- [26] P.A. Christenson, B.J. Willis, East Indian sandalwood oil. 1. Stereoselective synthesis of (+)-b-santalene and (+)-b-santalol, *J. Org. Chem.* 44 (1979) 2012-2018.
- [27] P.A. Christenson, B.J. Willis, East Indian sandalwood oil. 2. Stereoselective synthesis of (\pm)-*epi*- β -santalene and (\pm)-*epi*- β -santalol, *J. Org. Chem.* 45 (1980) 3068-3072.
- [28] W. Oppolzer, C. Chapuis, D. Dupuis, M. Guo, Asymmetric Diels–Alder reactions of neopentyl ether-shielded acrylates and allenic esters. Syntheses of (–)-norbornenone and (–)- β -santalene, *Helv. Chim. Acta* 68 (1985) 2100-2114.
- [29] S. Takano, K. Inomata, A. Kurotaki, T. Ohkawa, K. Ogasawara, Enantiodivergent route to both enantiomers of β -santalene and *epi*- β -santalene from a single chiral template, *J. Chem. Soc., Chem. Commun.* (1987) 1720-1722.
- [30] Y. Arai, M. Yamamoto, T. Koizumi, The asymmetric Diels-Alder cycloaddition using ethyl (–)Z-(R)S-2-methyl-3-(*p*-tolylsulfinyl)propenoate. Application to an enantioselective synthesis of (+)-*epi*-b-santalene, *Bull. Chem. Soc. Jpn.* 61 (1988) 467-473.
- [31] P.A. Unnikrishnan, P.A. Vatakencherry, Syntheses of *epi*- β -santalene, β -santalene and an isomer of β -santalene with 4-methyl-4-pentenyl side chain, *Synthetic Commun.* 22 (1992) 3159-3168.
- [32] T. Kamikubo, K. Ogasawara, Preparation of (+)-tricyclo[6.2.1.0_{2,7}]undec-2(7)-en-3-one and its conversion into (+)-*epi*-b-santalene, *Chem. Lett.* (1995) 95-96.
- [33] M. Saito, M. Kawamura, K. Ogasawara, Diastereo- and enantio-controlled synthesis of sandalwood constituents (–)- β -santalene and (+)-*epi*- β -santalene starting from the same (+)-norcamphor, *Tetrahedron Lett.* 36 (1995) 9003-9006.
- [34] C. Fehr, I. Magpantay, J. Arpagaus, X. Marquet, M. Vuagnoux, Enantioselective Synthesis of (–)- β -Santalol by a Copper-Catalyzed Enynol Cyclization-Fragmentation Reaction, *Angew. Chem. Int. Ed. Engl.* 48 (2009) 7221-7223.
- [35] W.I. Fanta, W.F. Erman, Novel use of borate protective groups in organic synthesis. A facile synthesis of dihydro- β -santalol, *Tetrahedron Lett.* (1969) 4155-4158.
- [36] W.I. Fanta, W.F. Erman, Synthesis of (\pm)-dihydro- β -santalol, *J. Org. Chem.* 37 (1972) 1624-1630.
- [37] J.A. Bajgrowicz, G. Frater, Chiral recognition of sandalwood odorants, *Enantiomer* 5 (2000) 225-234.
- [38] A. Kovatcheva, G. Buchbauer, A. Golbraikh, P. Wolschann, QSAR modeling of a-campholenic derivatives with sandalwood odor, *J. Chem. Inf. Comput. Sci.* 43 (2003) 259-266.
- [39] A.S. Dimoglo, A.A. Beda, N.M. Shvets, M.Y. Gorbachov, L.A. Kheifits, I.S. Aulchenko, Investigation of the relationship between sandalwood odor and chemical structure: electron-topological approach, *New J. Chem.* 19 (1995) 149-154.
- [40] R.E. Naipawer, K.L. Purzycki, G.W. Shaffer, R.E. Erickson In: **Essential Oils**, Mookherjee BD, Allured Publ. Corp. edn. Wheaton; 1981.
- [41] M. Chastrette, D. Zakarya, C. Pierre, Structure-odor relationships in sandalwood: search for an interaction model based on the concept of a santalophore superpattern, *Eur. J. Med. Chem.* 25 (1990) 433-440.
- [42] G. Buchbauer, A. Hillisch, K. Mraz, P. Wolschann, Conformational parameters of the sandalwood-odor activity: conformational calculations on sandalwood odor. Part X, *Helv. Chim. Acta* 77 (1994) 2286-2296.
- [43] B.D. Mookherjee: **New insights in the three most important natural fragrance products : wood, amber, and musk.** In: *Proceedings of the 12th International Congress on Flavour, Fragrance, and Essential Oils: 4-8 october 1992 1992; Vienna, Austria.* Austrian Association of Flavour and Fragrance Industry: 234-262.
- [44] E.J. Brunke, G. Schmaus, Les nouveaux composants de l'essence de bois de santal actifs sur le plan olfactif, 1^{ère} partie. Isolation et explication de la structure du cyclosantalal et de l'épi-cyclosantalal, *Dragoco Rep.* 42 (1995) 195.

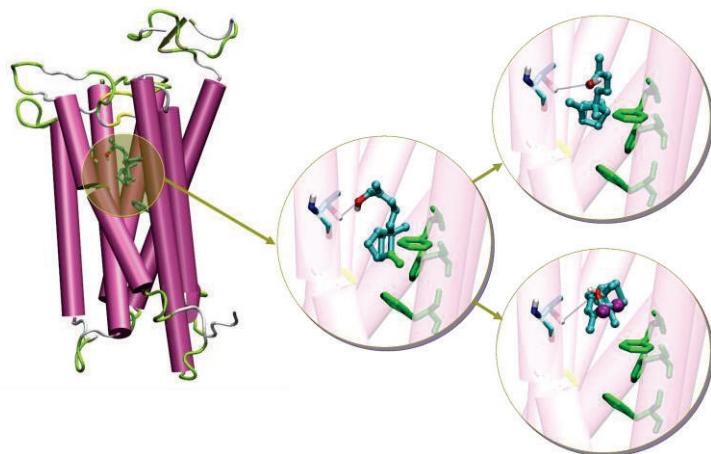
- [45] E.J. Brunke, E. Klein, in: E. Theimer (Eds.), *Fragrance Chemistry, The science of the sense of smell, Chemistry of Sandalwood Fragrance*, New York, 1982, pp. 397-431.
- [46] D.K. Kettenes, W. Lenselink: **Carane derivatives**. In: *German Patent*. DE: (Polak's Frutal Works B. V., Neth.). 1978: 49.
- [47] J.A. Bajgrowicz, P. Kraft: **Preparation of trimethylcyclopentenylalkoxyalkanols and -alkanals as odorous substances**. In: *Worldwide Patent*. WO: (Givaudan SA); 2008: 21pp.
- [48] C. Chapuis, P.-A. Blanc: **Campholenaldehyde derivatives and their use in perfumery**. In: *European Patent*. EP: (Firmenich S. A.); 1996: 20 pp.
- [49] T. Markert, V. Porrmann: **Preparation of (2,2,3-trimethyl-3-cyclopentenyl)allyl alcohols as perfume fragrances**. In: *Worldwide Patent*. WO: (Henkel K.-G.a.A.); 1993: 20 pp.
- [50] G. Buchbauer, F. Zechmeister-Machhart, P. Weiss-Greiler, P. Wolschann, Synthesis in the isocamphane series. Part 40. Structure-activity relationships of sandalwood odorants. *Synthesis and odor of methyl- β -santalol*, Arch. Pharm. 330 (1997) 112-114.
- [51] E.J. Corey, J.W. Suggs, Pyridinium chlorochromate. Efficient reagent for oxidation of primary and secondary alcohols to carbonyl compounds, *Tetrahedron Lett.* (1975) 2647-2650.
- [52] E. Demole, C. Demole, P. Enggist, A chemical investigation of the volatile constituents of East Indian sandalwood oil (*Santalum album L.*), *Helv. Chim. Acta* 59 (1976) 737-747.
- [53] T. Hasegawa, T. Toriyama, N. Ohshima, Y. Tajima, I. Mimura, K. Hirota, Y. Nagasaki, H. Yamada, Isolation of New Constituents with a Formyl Group from the Heartwood of *Santalum album L.*, *Flavour Fragr. J.* 26 (2011) 98-100.
- [54] E.J. Brunke, Nouveaux alcools sesquiterpéniques d'essence de bois de santal des Indes orientales, *Dragoco Rep.* 28 (1981) 91.
- [55] J.J. Brophy, C.J.R. Fookes, E.V. Lassak, Constituents of *Santalum spicatum* (R.Br.) A. DC. wood oil, *J. Essent. Oil Res.* 3 (1991) 381-385.
- [56] E.J. Brunke, Les produits aromatiques de santal. Structure chimique et note bois de santal, *Dragoco Rep.* (1981) 251.
- [57] E.J. Brunke, Le $(-)(Z)$ - α -trans-bergamotol, un nouveau constituant de l'essence de bois de santal des Indes orientales, intéressant sur le plan olfactif, *Dragoco Rep.* (1983) 27-32.
- [58] P. Weyerstahl, Odor and structure, *Journal fuer Praktische Chemie/Chemiker-Zeitung* 336 (1994) 95-109.
- [59] F. Ullrich, W. Grosch, Identification of the most intense volatile flavor compounds formed during autoxidation of linoleic acid, *Z. Lebensm.-Unters. Forsch.* 184 (1987) 277-282.
- [60] P. Weyerstahl, H. Marschall, U. Splittgerber, D. Wolf, H. Surburg, Constituents of Haitian vetiver oil, *Flavour and Fragrance J.* 15 (2000) 395-412.
- [61] M. Chastrette, Trends in structure-odor relationships, *SAR QSAR Environ. Res.* 6 (1997) 215-254.
- [62] S. Bieri, K. Monastyrskaya, B. Schilling, Olfactory Receptor Neuron Profiling using Sandalwood Odorants, *Chemical Senses* 29 (2004) 483-487.
- [63] B. Auger, J.A. Bajgrowicz, E. Giraudi: **Preparation of formylpinanes for fragrances**. In: *Worldwide Patent*. WO: Givaudan-Roure; 1993: 24 pp.
- [64] C. Chapuis: **Preparation of 3,3-dimethyl-5-(2,2,3-trimethyl-3-cyclohexenyl)-4-penten-2-ol enantiomers and analogs as perfume fragrances**. In: *European Patent*. EP: Firmenich; 1993: 16 pp.
- [65] I. Stappen, J. Hoefinghoff, S. Friedl, C. Pammer, P. Wolschann, G. Buchbauer, Structure-activity relationships of sandalwood odorants: Total synthesis and fragrance properties of cyclopropano- β -santalol, *Eur. J. Med. Chem.* 43 (2008) 1525-1529.
- [66] G. Buchbauer, P. Lebada, H. Spreitzer, P. Wolschann, Structure-odor relationships of sandalwood odorants: synthesis of (Z) -7-oxa- β -santalol, *Liebigs Ann. Chem.* (1995) 1693-1696.
- [67] G. Buchbauer, H. Spreitzer, F. Zechmeister-Machhart, A. Klinsky, P. Weiss-Greiler, P. Wolschann, Synthesis and olfactory activity of keto- β -santalol and methoxy- β -santalol, *Eur. J. Med. Chem.* 33 (1998) 463-470.

- [68] T. Aida, M. Harada, H. Iwai, A. Amano, T. Yamasaki, T. Yamamoto: **(E)-(R)-2-alkyl-4-(2,2,3-trimethylcyclopent-3-en-1-yl)-2-buten-1-ol, process for preparing the same, and use thereof in perfume compositions.** In: *European Patent*. EP: Takasago; 1998: 24 pp.
- [69] J.A. Bajgrowicz, G. Frater: **Preparation of optically pure isomers of campholenic aldehyde derivatives for use as detergent fragrances.** In: *European Patent*. EP: Givaudan-Roure; 1998: 11 pp.
- [70] C. Chapuis, A. Gautier, P.-A. Blanc: **Use of optically active isomers of (e)-3,3-dimethyl-5-(2,2,3-trimethyl-3-cyclopenten-1-yl)-4-penten-2-ol in perfumery.** In: *European Patent*. EP: Firmenich; 1995: 14 pp.
- [71] J.A. Bajgrowicz, G. Frater: **Preparation of cyclopentanebutanol-derivative odorants for perfumes.** In: *European Patent*. EP: Givaudan-Roure; 1997: 23 pp.
- [72] J.A. Bajgrowicz, I. Frank, G. Frater, M. Hennig, Synthesis and structure elucidation of a new potent sandalwood-oil substitute, *Helv. Chim. Acta* 81 (1998) 1349-1358.
- [73] J.M. Castro, P.J. Linares-Palomino, S. Salido, J. Altarejos, M. Nogueras, A. Sanchez, Enantiospecific synthesis, separation and olfactory evaluation of all diastereomers of a homologue of the sandalwood odorant Polysantol, *Tetrahedron* 61 (2005) 11192-11203.
- [74] B. Hoelscher, N.A. Braun, B. Weber, C.-H. Kappey, M. Meier, W. Pickenhagen, Enantioselectivity in odor perception synthesis and olfactory properties of the new tricyclic sandalwood odorant Fleursandol, *Helv. Chim. Acta* 87 (2004) 1666-1680.
- [75] J.G. Witteveen, A.J.A. Van der Weerdt, Structure-odor relationships of some new synthetic sandalwood aroma chemicals. Synthesis and olfactive properties in a series of bicyclo[4.4.0]decan-3-ols, *Recueil des Travaux Chimiques des Pays-Bas* 106 (1987) 29-34.
- [76] G. Buchbauer, H. Spreitzer, H. Swatonek, P. Wolschann, Absolute configuration and odor analysis of the enantiomeric tert-butylbicyclo[4.4.0]decan-3-ols, *Tetrahedron-Asymmetr.* 3 (1992) 197-198.
- [77] K.H. Shankaranarayana, K. Parthasarathi, Synthetic sandalwood aroma chemicals, *Perfum. Flavor.* 9 (1984) 17-20.

Article 6:

An Olfactophore model predicts Odorant Receptors involved in β -Santalol discrimination

Claire A. de March, Céline Delasalle, SangEun Ryu, Fouzia El Mountassir, Anne-Marie Le Bon, Cheil Moon, Nicolas Baldovini, Jérôme Golebiowski, en préparation.



Keywords: β -santalol, Olfactory receptor, Molecular modeling, Olfactophore

Abstract

The identification of chemogenomic links connecting chemical and odorant spaces to that of odorant receptors would allow a rational design of odorant compounds. Using the sandalwood olfactophore model, we identified the features required for an odorant receptor for differentially responding to β -santalol, its dimethyl derivative or to α -santalol. Molecular modeling provides a rationale for the discrimination of these three odorants in hOR1G1, as assessed *in vitro*. This provides a new working hypothesis where the information on receptors involved in odor coding is hidden within olfactophores.

Introduction

The strategy used by our brain to detect airborne chemicals relies on a combinatorial code of Odorant Receptors (OR) expressed by our Olfactory Sensory Neurons (OSN).[1] OR are indeed the cornerstone of our sense of smell, as they allow our brain to evaluate the nature, the concentration and the origin of chemicals present in our environment. Our genome contains ~1000 OR genes, of which 396 are considered functional.[2] This large number of functional OR endows us with an extraordinary discriminating power.[3] Since their discovery, many studies have been published with the purpose of their deorphanization. To date, 57 humans ORs have been deorphanized with various sets of molecules differing in shapes or chemical functions.[4]

However, even using elaborated chemometrics protocol aiming at covering as widely as possible the chemical space, screening of responding ORs with chemicals supposed to cover this odorant space was not easy. Data-mining protocols were developed to tentatively identify chemo-genomic links, on the basis of physicochemical descriptors of odorants, proposing a way to accurately select odorant for screening receptors.[5] A similar protocol proved particularly efficient for deorphanizing insect ORs.[6, 7]

Screening of hundreds of OR with more than 60 odorants produced a hit rate of less than 6% (27 ORs showed a significant response amongst 511 tested), suggesting that the odorant space is actually very subtle and that chemogenomic links are particularly tricky for odorant receptors.[8, 9]

In a more direct approach, chemists tried for a long time to connect the chemical structures of odorants with the odor they are associated with. These structure-odor relationships (SOR) connect common chemical features of odorants with prototypical odors.[10, 11] Getting inspired from the pharmacophore approach, the so-called Olfactophore recapitulates common physico-chemical features within a series of chemicals associated to the same odor family and proceeds to their optimal superimposition in space.[11] To what extend can one infer information on the receptors

involved in the combinatorial code of a given smell? It is very likely that information about OR is hidden within these olfactophore models.

Here, we provide a proof of principle for the identification of a receptor space associated to a chemical space through an analysis of an olfactophore and its comparison with a receptor pharmacophore. A comparison between the two properties suggests that hOR1G1 is involved in the discrimination of β -santalol from its derivatives. Cellular functional tests and molecular modeling studies were performed on the receptor bound to β -santalol, α -santalol and dimethyl β -santalol. β -santalol activates hOR1G1 while α -santalol and its (odorless) dimethyl analogue seems are much less associated to receptor activation. The physico-chemical features within the olfactophore and the pharmacophore are associated to the presence of specific amino-acids within the receptor binding cavity and their interaction with the three odorants.

Results and discussions

The olfactophore of sandalwood odor have been made up by comparing common features of sandalwood and non-sandalwood odorants.[12] This brought us to ask which ORs are responsible for this perception. Although this question will probably remain unanswered for a time, we noticed that the human OR1G1 was activated by long aliphatic odorants as well as by camphor, which share chemical features with the santalane family.[13-15] Intriguingly, the santalane olfactophore hypothesis is quite identical to that of the hOR1G1 pharmacophore obtained with the set of odorants used to deorphanize this OR, published by Sanz et al.[14] Figure 1 compares the two models, which both show a hydrogen-bond acceptor and two hydrophobic sites. The Root Mean Square deviation between the two models of 0.3 Å suggests that hOR1G1 is involved in the differential perception of β -santalol with respect to other related odorant structures.

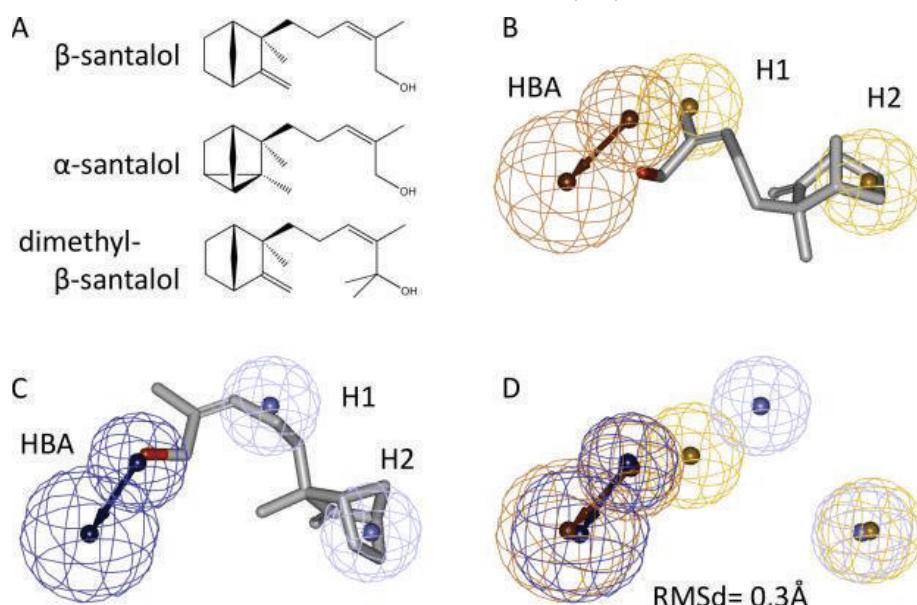


Figure 1. Chemical structures of β -santalol and its derivatives (A). Tridimensional models of pharmacophore of hOR1G1 (B) and sandalwood olfactophore (C). In each case, β -santalol is shown in its optimal fit to the model. H-bond acceptor features are represented by two spheres (HBA: brown and dark blue in B and C, respectively) and hydrophobic features are shown by single spheres (H1-H2: yellow and light blue in B and C, respectively). Their superimposition (D) presents a RMSd of 0.3 Å.

To assess that OR1G1 is a receptor involved in the discrimination between some santalane derivatives, we have performed calcium imaging assays on three selected components, to evaluate their ability to activate this receptor.

β -santalol was chosen as the prototypical sandalwood odorant bearing a santalane skeleton. Compared to β -santalol, α -santalol presents a subtle modification at the polycyclic ring, but quite different olfactory properties. These odorants were clearly separated by the olfactophore model.[12] Another compound was also evaluated: the odorless dimethyl- β -santalol which is also well separated from β -santalol by the olfactophore model.

hOR1G1 has been heterogeneously expressed in Hana3A cells and its activation by our three selected molecules was monitored by single-cell calcium imaging. Tridecanal, a known strong agonist is measured as a reference.[14] The results are reported in Figure 2.

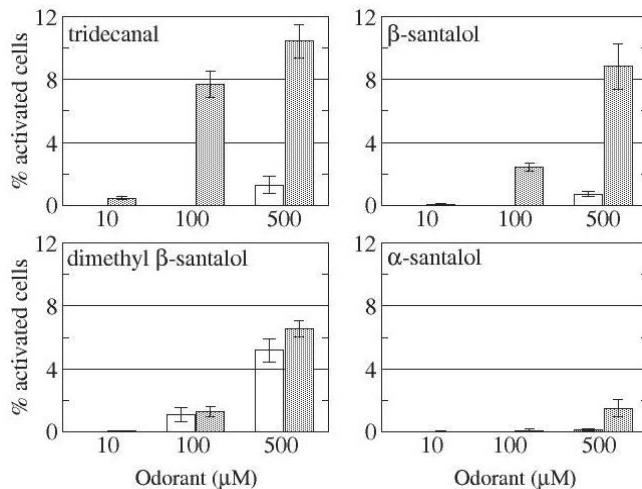


Figure 2. Response of Hana3A cells with either no receptor (white bars) or expressing hOR1G1 (grey bars). Odorants are compared to tridecanal.

We measure a differential activation of hOR1G1 depending on the studied odorant. β -santalol behaves as a strong agonist, with a dose-dependent activation. α -santalol does not elicit any sufficient dose-response activation of the cells below a concentration of 500 μ M, where a low activation is observed revealing that α -santalol is a weak agonist of hOR1G1 *in vitro*. Dimethyl- β -santalol has a similar behavior, as it only activate cells expressing hOR1G1 at high concentration. These results suggest that hOR1G1 is involved in the combinatorial code of β -santalol perception. This receptor actually contributes to the odor discrimination between β -santalol and the structurally related odorants α -santalol and dimethyl- β -santalol.

As a matter of fact, the olfactophore approach seems to embed information about the nature of the olfactory receptors involved in the combinatorial code of the perception of smell. To go further into the understanding of the chemo-genomic links that makes our sensory system so subtle, we have used a molecular model of hOR1G1 to build the complexes with the odorants studied above.

The differential binding is studied at the atomic-level through a molecular modeling approach. The structure of hOR1G1 is experimentally unknown but we have built a model of hOR1G1[16] in line with available site-directed mutagenesis data.[17] The structure of hOR1G1 is built by homology modeling with available X-Ray data and ligands are further docked into the binding site.[18]

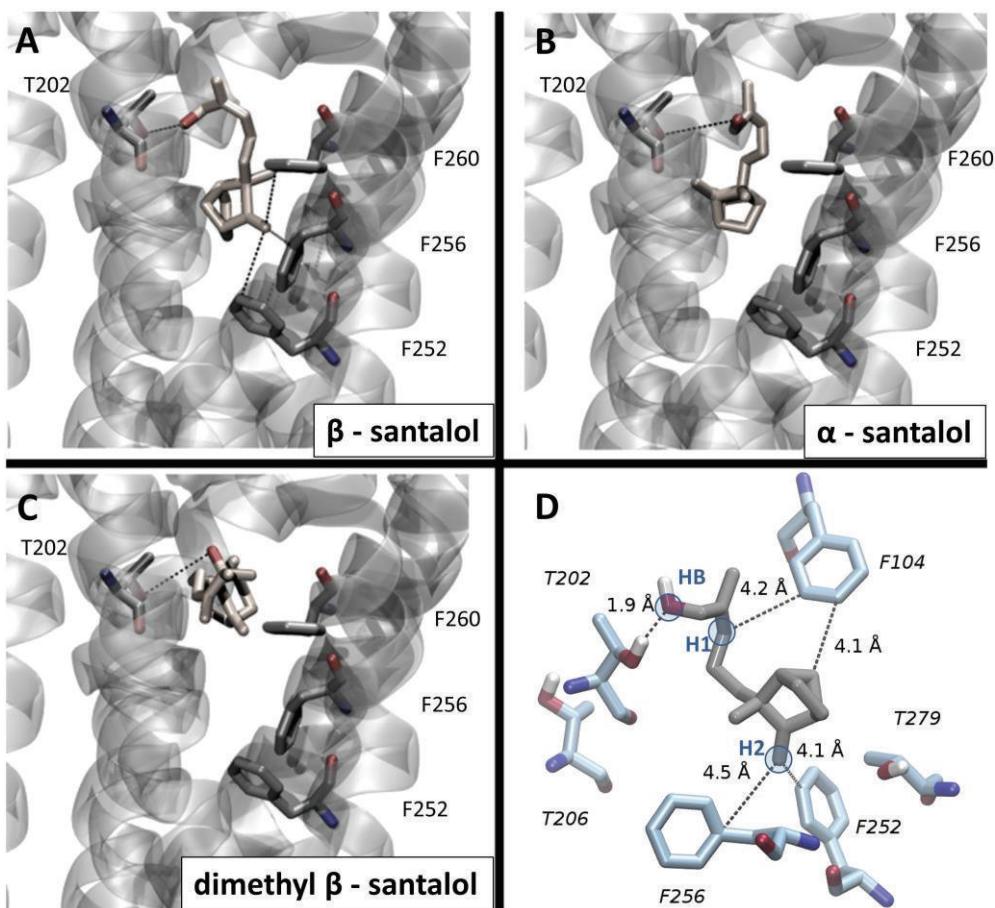


Figure 3. Models of the three odorants (A, B, and C) bound into the cavity of hOR1G1. hOR1G1 secondary structure is shown as a grey cartoon. Odorants and residues T202, F252, 256 and 260 are shown explicitly. Hydrogen atoms are omitted for clarity. The colors are as follows: C atoms of the ligands: pink, of the receptor: grey; O atoms: red, N: blue. A: β -santalol shows two interactions, one with TM5 (T202) and the other with several residues in TM6 (F252, F256, and F260). B: α -santalol only engages an interaction with T202. C: similarly, dimethyl β -santalol only interacts with T202. D: detailed view of β -santalol bond to hOR1G1. F104, T202, T206, F252, 256 and T279 are shown explicitly and hydrogen atoms are omitted for clarity. The residues within hOR1G1 binding cavity are consistent with the features of the olfactophore/pharmacophore: typical physico-chemical features are symbolized by blue circles (HB corresponds to H-bond feature and H1 and H2 to hydrophobic features).

From a general point of view, the anchoring point of these odorants to hOR1G1 cavity (T202, *vide infra*) is conserved but the remaining interactions are different. The structure of the β -santalol-hOR1G1 complex shows an optimal interaction between the various parts of the odorant and several receptor residues. As shown in figure 3, the hydroxyl group of the ligand establishes a hydrogen bond with T202 residue in TM5. This residue is aligned with a serine residue (S203) in the β 2-adrenergic receptor, crucial for the recognition of its catecholamine agonists.[19] The bulky part of the ligand lies in the region of the binding site that gathers several aromatic residues of TM3 and TM6 (Phe 104,

252 and 256). Interestingly, the double bonds H1 and H2 of β -santalol are in contact with these aromatic residues, showing an optimal interaction between the π -electrons of both protagonists. Dimethyl- β -santalol differs only by the presence of two carbon atoms at the alcohol moiety. In the complex, the H-bond interaction with T202 is, although weaker, still observed (heavy atom distance of 4.4 Å, instead of 1.9 Å for β -santalol as shown in figure 3), but the presence of the two additional methyl groups precludes rotation of the rest of the structure interaction of the bulky part of the ligand with the aromatic residues of the OR pocket. Instead, this bulky group is now closer to Phe 104, at the opposite side of the pocket but remains too far to engage a strong π -interaction with it. α -santalol adopts a structure similar to that of β -santalol, but the lack of a π -bond at its bulky group abolishes the strong interaction with the aromatic residues.

The residues engaged in the recognition of β -santalol are highly variable within the OR family, suggesting that hOR1G1 can selectively recognize this odorant with respect to many other ORs. These results provide a rationale on the relation between the olfactophore features of the santalane hypothesis and the OR1G1 binding cavity. The Hydrogen Bond Acceptor area (HBA) of the olfactophore corresponds to an interaction with T202 through a hydrogen bond in OR1G1 (Figure 3). The activation of OR1G1 must be also associated to an interaction with a hydrophobic area corresponding to the Phenylalanine residues of the receptor. Of course, the toogle-switch of the receptor, Y252, is engaged in the interaction.[20] This interaction and the other hydrophobic contacts are modeled by the two hydrophobic features (H1 and H2) in the olfactophore (Figure 3). In addition, the H-bond interaction should not be hindered by a bulky vicinal group, which is associated to several exclusion volumes in the olfactophore.[12] Only β -santalol is able to satisfy all criteria (H-bond, Hydrophobic contacts, no atom within exclusion volume). This is consistent with the santalane hypothesis of our olfactophore approach, able to discriminate β -santalol from both α -santalol and dimethyl- β -santalol. This suggests that one of the sandalwood olfactophore hypotheses corresponds to the activation of ORs closely related to hOR1G1, where residues similar to the three phenylalanine residues (104, 252, 256 and 260) and the threonine (202) are conserved. Investigations on other receptors containing only parts of these residues in their sequence (hOR13H1 and hOR52A1) are in progress (Figure4). The receptor hOR13H1 exhibits the same HBA hydophobic properties compared with hOR1G1. It indeed has two threonines residues in TM5 at position 202 and 206 but it lacks hydrophobic features H1 and H2 in TM3 and TM6. Contrarily, hOR52A1 possesses the hydrophobic features (H1 and H2) with the presence of two valine residues in TM3 and three aromatic residues in TM6 but lacks the H-bond acceptor HBA with the presence of isoleucines instead of threonines in TM5.

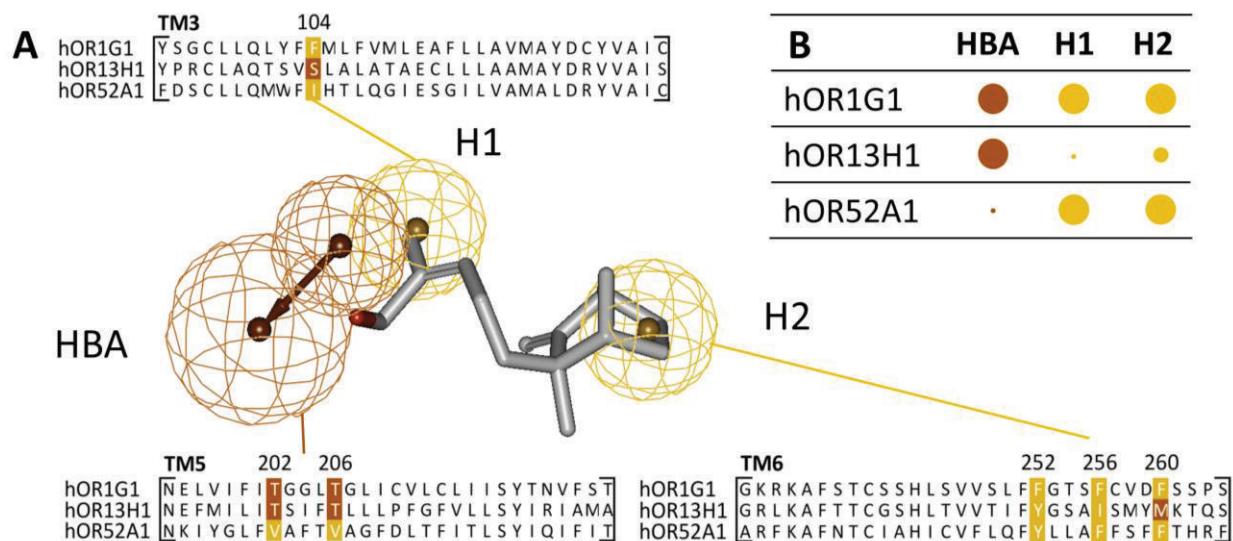


Figure 4. Comparison of binding cavity features of hOR1G1, hOR13H1, and hOR52A1. (A) Pharmacophore of hOR1G1 (center). The three different features (H1, H2 and HBA) are linked with different parts of the sequences (TM3, TM5 and TM6) of hOR1G1, hOR13H1 and hOR52A1, exhibiting differences at the binding cavity of these receptors. These differences are highlighted in a table (B). Large circles represent the presence of a feature in a receptor and a dot the lack of this feature. H-bond acceptor (HBA) properties are represented in brown and hydrophobic ones (H1 and H2) are shown in gold.

The bioinformatic analysis of the sequence of hOR1G1 in comparison with sequences of the all set of OR allows the rational identification and selection of ORs likely to respond differentially to the set of β -santalol derivatives. These two other receptors should interact differently with the set of odorants highlighting the importance of the three features within the receptor space in the recognition of this chemical space.

Conclusion

It is now known that the effect of an odorant molecule on the sensory system is the result of the activation of several olfactory receptors. These receptors can either have a large spectrum of recognition (as such, they can be considered as broadly-tuned) or be narrowly-tuned towards a given chemical family. Here, we have shown that although α -santalol, β -santalol and its tertiary alcohol derivative are very similar, they can nevertheless be discriminated by hOR1G1, a broadly-tuned receptor.[15]

The number of receptors involved in the sandalwood perception is unknown. They might however share some binding cavity features that will have to match with those found here. These features are in majority hydrophobic and fulfilled by few hydrophilic residues. The olfactophores found previously[12] can serve as a useful tool to perform virtual screening of ORs binding cavity when

Article 6 – de March et al. en préparation

these latter will be available either by means of molecular modeling methods or by experimental ones. For the moment, only few human OR structures have been modeled theoretically,[14, 15, 21-23] but newer one may be available in the next future to continue fulfilling the chemogenomic links between odorants and OR sequences. In addition, the data on a specific OR can be extrapolated on other receptors by means of bio-informatics analyses. It is a way to rationally select ORs of interest for a set of fragrance compounds.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] Y. Niimura, Olfactory Receptor Multigene Family in Vertebrates: From the Viewpoint of Evolutionary Genomics, *Current Genomics* 13 (2012) 103-114.
- [3] C. Bushdid, M.O. Magnasco, L.B. Vosshall, A. Keller, Humans Can Discriminate More than 1 Trillion Olfactory Stimuli, *Science* 343 (2014) 1370-1372.
- [4] C.A. de March, S. Ryu, G. Sicard, C. Moon, J. Golebiowski, Structure–odour relationships reviewed in the postgenomic era, *Flavour and Fragrance Journal*, 10.1002/ffj.3249(2015) n/a-n/a.
- [5] R. Haddad, R. Khan, Y.K. Takahashi, K. Mori, D. Harel, N. Sobel, A metric for odorant comparison, *Nat. Methods* 5 (2008) 425-429.
- [6] S.M. Boyle, S. McInally, A. Ray, L. Luo, Expanding the olfactory code by in silico decoding of odor-receptor chemical space, *eLife* 2 (2013) e01120.
- [7] P. Kain, S.M. Boyle, S.K. Tharadra, T. Guda, C. Pham, A. Dahanukar, A. Ray, Odour receptors and neurons for DEET and new insect repellents, *Nature* 502 (2013) 507-512.
- [8] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, J.D. Mainland, Odor Coding by Mammalian Receptor Repertoire, *Science Signaling* 2 (2009) ra9.
- [9] J.D. Mainland, A. Keller, Y.R. Li, T. Zhou, C. Trimmer, L.L. Snyder, A.H. Moberly, K.A. Adipietro, W.L.L. Liu, H. Zhuang *et al*, The missense of smell: functional variability in the human odorant receptor repertoire, *Nat Neurosci* 17 (2014) 114-120.
- [10] K.J. Rossiter, Structure-Odor Relationships, *Chem. Rev.* 96 (1996) 3201-3240.
- [11] P. Kraft, J.A. Bajgrowicz, C. Denis, G. Frater, Odds and trends: recent developments in the chemistry of odorants, *Angew. Chem. Int. Ed. Engl.* 39 (2000) 2980-3010.
- [12] C. Delasalle, C.A. de March, U.J. Meierhenrich, H. Brevard, J. Golebiowski, N. Baldovini, Structure-Odor Relationships of Semisynthetic β-Santalol Analogs, *Chemistry & Biodiversity* 11 (2014) 1843-1860.
- [13] G. Sanz, C. Schlegel, J.-C. Pernollet, L. Briand, Comparison of Odorant Specificity of Two Human Olfactory Receptors from Different Phylogenetic Classes and Evidence for Antagonism, *Chemical Senses* 30 (2005) 69-80.
- [14] G. Sanz, T. Thomas-Danguin, H. Hamdani el, C. Le Poupon, L. Briand, J. Pernollet, E. Guichard, A. Tromelin, Relationships between molecular structure and perceived odor quality of ligands for a human olfactory receptor., *Chem. Senses* 33 (2008) 639-653.
- [15] L. Charlier, J. Topin, C. Ronin, S.-K. Kim, W. Goddard, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cell Mol Life Sci* (2012) 4205-4213.
- [16] C.A. de March, S.-K. Kim, S. Antonczak, W.A. Goddard, J. Golebiowski, G protein-coupled odorant receptors: From sequence to structure, *Protein Sci*, 10.1002/pro.2717(2015).
- [17] G. Launay, S. Teletchea, F. Wade, E. Pajot-Augy, J.F. Gibrat, G. Sanz, Automatic modeling of mammalian olfactory receptors and docking of odorants, *Protein Engineering Design & Selection* 25 (2012) 377-386.
- [18] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, in: C.J. Crasto (Eds.), *Olfactory Receptors, Molecular Modelling of Odorant/Olfactory Receptor Complexes*, 2013, pp. 53-65.
- [19] A.M. Ring, A. Manglik, A.C. Kruse, M.D. Enos, W.I. Weis, K.C. Garcia, B.K. Kobilka, Adrenaline-activated structure of beta(2)-adrenoceptor stabilized by an engineered nanobody, *Nature* 502 (2013) 575-579.

Article 6 – de March et al. en préparation

- [20] C.A. de March, Y. Yu, M.J. Ni, K.A. Adipietro, H. Matsunami, M. Ma, J. Golebiowski, Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors, *Journal of the American Chemical Society* 137 (2015) 8611-8616.
- [21] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *J Struct Biol* 159 (2007) 400-412.
- [22] L. Doszczak, P. Kraft, H.-P. Weber, R. Bertermann, A. Triller, H. Hatt, R. Tacke, Prediction of Perception: Probing the hOR17-4 Olfactory Receptor Model with Silicon Analogues of Bourgeonal and Lilial, *Angewandte Chemie International Edition* 46 (2007) 3367-3371.
- [23] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angewandte Chemie International Edition* 51 (2012) 1274-1278.

Article 7 - Le calcul de l'affinité odorant/récepteur discrimine les agonistes des non-agonistes de hOR1G1.

La génération du signal olfactif depuis la molécule odorante jusqu'à la perception consciente passe par l'interaction de cette molécule avec nos récepteurs olfactifs. L'identification des couples odorant/RO(s) est donc nécessaire au décryptage du code combinatoire associé aux odeurs.

Le récepteur olfactif va déclencher un signal si le ligand est capable de remplir deux conditions : (1) avoir une certaine affinité avec le site de liaison du récepteur, (2) déclencher des changements conformationnels du RO pour engendrer son couplage à la protéine G, à l'origine de la signalisation cellulaire (Figure 1). Ces deux étapes peuvent être associées à des équilibres chimiques schématisés dans la figure 1.

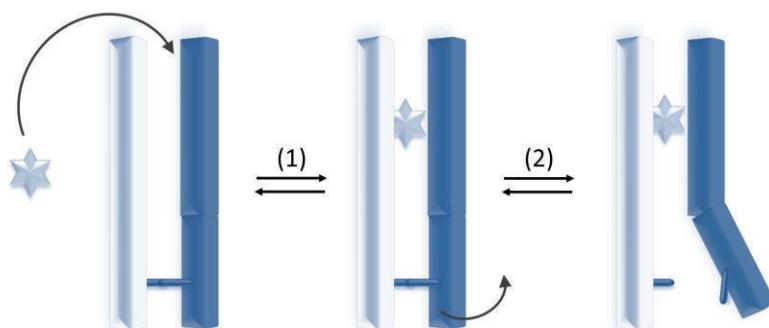


Figure 1. La voie de signalisation typique des RCPGs est associée à deux équilibres consécutifs. L'équilibre (1) correspond à l'affinité entre le ligand (étoile) et la cavité de liaison du RO. L'équilibre (2) est associé à un changement conformationnel du RO qui engendre une ouverture au niveau de la partie intracellulaire favorisant le couplage à la protéine G.

La méthode privilégiée ici pour déterminer la capacité d'une molécule odorante à remplir ces deux critères est la modélisation moléculaire. L'affinité d'un ligand avec la cavité d'un récepteur peut être déterminée grâce au calcul de la variation de l'enthalpie libre de liaison ($\Delta G_{\text{liaison}}$). Dans cet article, cette première étape est privilégiée, en utilisant le calcul d'affinité pour discriminer les agonistes des non-agonistes d'un RO à large spectre (hOR1G1). Nous montrons que cette affinité est un descripteur pertinent dans le cas d'un groupe de dix molécules (huit agonistes, deux non-agonistes) et que les résultats de modélisation moléculaire prédisent les données expérimentales obtenues par imagerie calcique *in vitro*.

PARTIE 3 : Relations structure-fonction des récepteurs olfactifs

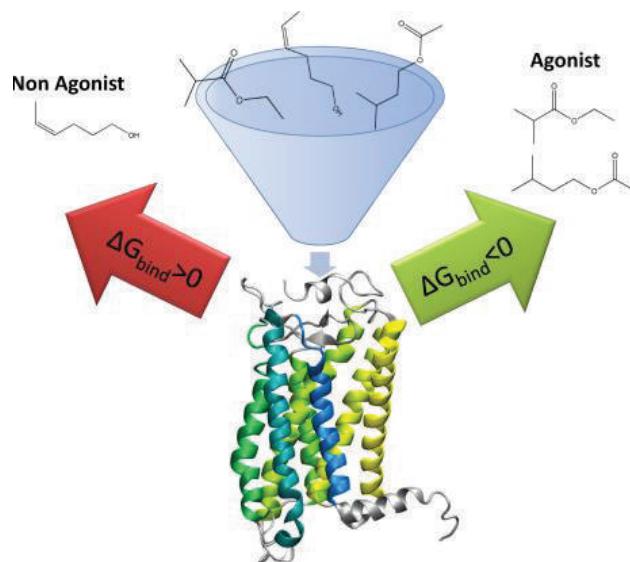
Article 7 – Topin, de March et al. Chem. Eur. J. 20 (2014) 10227-30

Pour cet article, j'ai construit les structures initiales pour les simulations de dynamique moléculaire en réalisant les calculs de docking des ligands. J'ai également construit un modèle alternatif du récepteur olfactif 1G1 afin de créer une discussion critique sur la pertinence du modèle et de son protocole de reconstruction.

Article 7:

Discrimination between Olfactory Receptor agonists and non-agonists

Jérémie Topin, Claire A. de March, Landry Charlier, Catherine Ronin, Serge Antonczak and Jérôme Golebiowski, Chemistry – A European Journal, 20 (2014) 10227–30

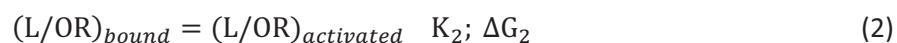
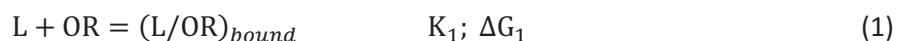


Keywords: odorant, receptor, free energy, molecular modeling, calcium imaging.

Abstract

We report on a joint approach combining free energy calculations and calcium imaging assays on the broadly-tuned human 1G1 Olfactory Receptor. The free energy of binding of ten odorants is computed by means of molecular dynamics simulations. This state function allows separating the experimentally determined eight agonists from the two non-agonists. This study constitutes a proof of principle for the computational deorphanisation of Olfactory Receptors.

Our nervous system is notably in charge of computing the external signals coming from its environment. Amongst these exogenous signals, odors are arguably the most complex, due to the large chemical space they are associated to. At the molecular level, each Olfactory Receptor Neuron (ORN) expresses only one type of Olfactory Receptor (OR) that acts as the cornerstone of the perception.[1, 2] The nature of these receptors has been evidenced by the pioneering work of Buck and Axel.[3] A single ligand can elicit a response by several of the 396 functional ORs in human beings, endowing us an extraordinary discriminating power.[4, 5] The originality of the perception of smell stems from the fact that odorants cover all chemical families and can vary in size, shape and even stereochemistry. Therefore it is virtually unrealistic to experimentally screen all ORs with all odorants. Although several atomic-level studies have used molecular modeling, they were mostly focused on gaining structural information on the complexes between ORs and odorants or identifying residues that control receptor's selectivity, (see ref.[6] for a review). ORs belong to the family of class-A G-protein Coupled Receptors and as such, their activation mechanism[7] can be split into two main events. First, the free odorant and the free OR are in a chemical equilibrium with a bound state (eq. 1). If the ligand is thermodynamically favored in the bound state, the complex will be the subject of a second equilibrium. In the case of an agonist, the active state of the complex will be favored. In presence of an antagonist, the receptor will remain in its inactive form (eq. 2).

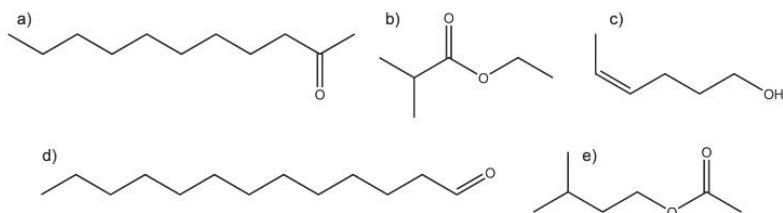


In this article, we report on an approach that focuses on the first equilibrium by using state-of-the art molecular simulation protocols after having experimentally assessed the odorants potency. The latter is compared to the odorant free energy of binding with the receptor, computing ΔG_1 of eq. 1 as an ensemble calculation of binding energy using averages accrued from long scale molecular dynamics (eq. 3).

$$\Delta G_1 = \Delta G_{binding} = \langle \Delta G_{complex} \rangle - \langle \Delta G_{OR} \rangle - \langle \Delta G_{odorant} \rangle \quad (3)$$

An equivalent approach based-on docking scores was reported in 2005.[8] We have however previously shown that a docking score cannot be used as a totally predictive value to discriminate between agonists and non-agonists of the broadly-tuned hOR1G1 receptor.[9]

The ligands considered here are shown in Scheme 1. They fulfill a previous series made up of nonanal, nonanol, 9-decen-1-ol, camphor and n-butanal.[9]



Scheme 1. Structure of the studied ligands: a) 2-undecanone, b) ethyl isobutyrate, c) cis-4-hexen-1-ol, d) tridecanal and e) isoamyl acetate.

Figure 1 reports fura-2 fluorescence Ca^{2+} imaging assays on four ligands with hOR1G1. The data result from 72H post-infected Sf9 cells and are compared to that of an endogenous receptor stimulated by octopamine at 50 μM .[10] A dose-response analysis was performed beforehand for isoamyl acetate (fig. 1, insert). The Sf9 cells respond strongly to the odorant at the three concentrations considered (25, 50 and 100 μM) without inducing cell toxicity. In the three cases, the high potency of the odorant is put forward. Accordingly, the assays are performed with the other odorants at concentrations of 100 μM .

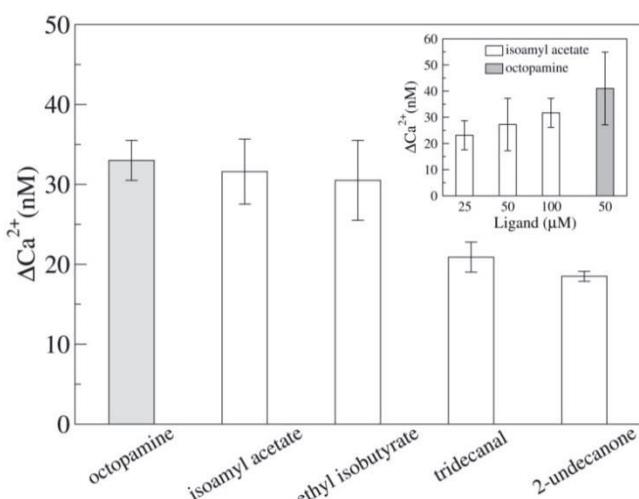


Figure 1. hOR1G1 calcium imaging assays results, shown as ΔCa^{2+} discharge for each odorant, compared to the average response triggered by octopamine at 50 μM . SEM (Standard Error of the Mean) is shown as error bars. Cis-4-hexenol does not activate hOR1G1, as shown in ref.[11].

The resulting calcium discharges are comparable to that of octopamine, assessing that 2-undecanone, ethyl isobutyrate, tridecanal and isoamyl acetate are strong agonists of hOR1G1. They induce similar intracellular calcium concentrations to other agonists measured with the same protocol emphasizing the broadly-tuned character of this OR.[9, 12] 2-undecanone is the agonist that triggers the weakest calcium discharge but the latter remains more than half of that of octopamine or isoamyl-acetate. These five systems are studied at the atomic-level by means of a multiple Molecular Dynamics (MMD) protocol.

The structure of hOR1G1 was build using Rhodopsin as a template following the protocol of Charlier et al. [9, 13] Opsin was shown recently to be a pertinent model for building OR structures.[14-16] Several studies on ORs considering the same template were shown to be able to reproduce the effect of site-directed mutagenesis on odorant activity, suggesting that OR cavities were accurately modelled with such a template.[6] Lai et al performed a thorough sequence analysis between an OR and different templates and confirmed that the use of rhodopsin is good starting point for homology modelling.[17] Figure 2 shows the structure of hOR1G1 and its binding cavity. Notice that some amino-acids (112, 181, 206, 252, 279) were proposed as belonging to the binding site based-on a residue conservation analysis of ORs.[18] Other residues of our model were however not predicted as belonging to the binding site (202, 259, 256, 260) emphasizing that atomic-level models capture the ligand-receptor interface with a higher accuracy. Generally, our model is in line with available site-directed mutagenesis on Ala 112 residue.[19] This amino acid is indeed located at the bottom of the binding pocket and its mutation affects odorant recognition.

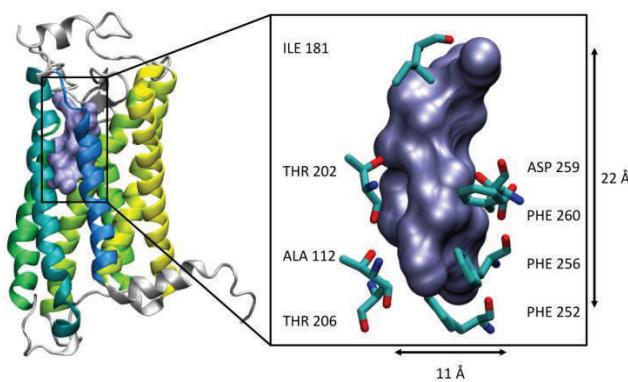


Figure 2. hOR1G1 model. The seven transmembrane helices are colored from yellow to blue. On the right, a close view of the binding cavity (calculated using Fpocket[20, 21]) with amino acids lining the cavity surface.

As observed in two independent previous studies, van der Waals interactions are the major contributor to binding, accounting for more than 60% of the binding energy.[9, 22] A per residue decomposition of the binding energy is provided in Table S4.

Most docking poses exhibit contact between the odorant and Thr 202 and Thr 206. These threonine residues are highly variable in human OR sequences suggesting that they actively participate to the differential recognition spectrum between all human ORs. Notice that these two residues are aligned with two serine residues in the β 2-adrenergic receptor, which act as the main anchoring points for the catecholamine agonists.[23] Thr 279 was also shown to be involved in recognition of hOR2AG1 agonists.[14] Docking poses involving directions towards these residues were selected for further Molecular Dynamics (MD) simulations. The hydrogen bonds observed in the docking poses are rapidly broken and their contribution is deemed rather minor (see Table S5).

To sample accurately the conformational space covered by the systems, two independent 100 ns MD simulations for each ligand are performed. They differ from the initial pose of the odorant within the binding site. These poses are chosen with totally opposite direction of the hydrophilic part of the ligand (toward either TM3 or TM7) to allow sampling of a large conformational space during the two MDs (see Table S1). The last 50 ns of these simulations were subjected to a MM-GBSA free energy component calculation associated to a normal mode analysis for estimating the entropy component change upon binding. The final score is the average of the data obtained from the 2 simulations. Table 1 gathers the energy analysis.

The free energy of binding ranges between -12.3 and +1.0 kcal·mol⁻¹ for tridecanal and cis-4-hexenol, respectively. Interestingly, the methods recover that all agonists show a negative (stabilizing) free energy of binding while cis-4-hexenol has a non-binder behavior, with a positive ΔG_{bind} . The free energy component associated to the internal energy and solvation is systematically negative. This term is partly compensated by the vibrational entropic contribution. In the case of cis-4-hexenol, this unfavorable contribution overcompensates the stabilization term, leading to an average positive free energy of binding. The agonists still show stabilizing free energies of binding after this term is accounted for.

Table 1. Free energy analysis for the five odorants bound to hOR1G1. For each odorant, two simulations are analysed and averaged. See methods for explanations of the different terms. Values are in kcal·mol⁻¹. SEM is shown within parenthesis. The experimental potency is indicated.

Odorant	$\Delta G_{\text{MM-GBSA}}$	T. $\Delta S_{\text{vib.}}$	$\langle \Delta G_{\text{bind}} \rangle$	potency ^a
tridecanal	-31.8 (0.1)	-19.5 (0.8)	-12.3 (0.6)	*
2-undecanone	-28.0 (0.2)	-20.5 (0.7)	-7.5 (0.9)	*
ethyl isobutyrate	-18.9 (0.6)	-13.6 (0.9)	-5.3 (1.4)	*
isoamyl acetate	-18.4 (1.4)	-15.9 (0.8)	-2.5 (0.6)	*
cis-4-hexenol	-17.5 (5.0)	-18.5 (0.6)	1.0 (4.4)	-

a: * means agonist; - means non-agonist

In the case of cis-4-hexenol, the two calculated free energies of binding largely differ. This result highlights two different behaviors of the cis-4-hexenol according to the starting pose of the simulation. In one simulation, associated to a weak negative free energy of binding (-3.4 kcal·mol⁻¹), the ligand occupies the upper part of the cavity. It is stabilized by a hydrogen-bond with Ile181 backbone from the Extra-Cellular Loop 2. This position is similar to a transient position observed in complexes involving the β 2-adrenergic receptor but is not meant to activate the receptor.[24] During the second simulation, the ligand is located deeper within the binding cavity as observed for other odorants. This pose is however associated to a positive free energy of binding (+5.4 kcal·mol⁻¹, see supporting information) and the resulting average between the two simulations is positive. This underlines the strength of a MMD protocol, which, by sampling different configurations of the system, is able to discriminate between our four binders and the non-binder.

Agonists show rather different behavior with respect to cis-4-hexenol. Each of them samples equivalent positions within the binding cavity during the two simulations as emphasized by the small standard error associated to the average free energy of binding in Table 1. They all have negative free energies of binding.

To fulfill the set of these five ligands, we have revisited molecular mechanics and solvation free energy components of known agonists and non-agonists of hOR1G1 by computing their vibrational entropic component to binding.[9] This allows gaining information on a more diverse set of odorants for this receptor. The set now contains eight agonists of variable lengths, structures and chemical functions and two non-agonists also belonging to two different chemical classes. Figure 3 gathers all free energy of binding calculations.

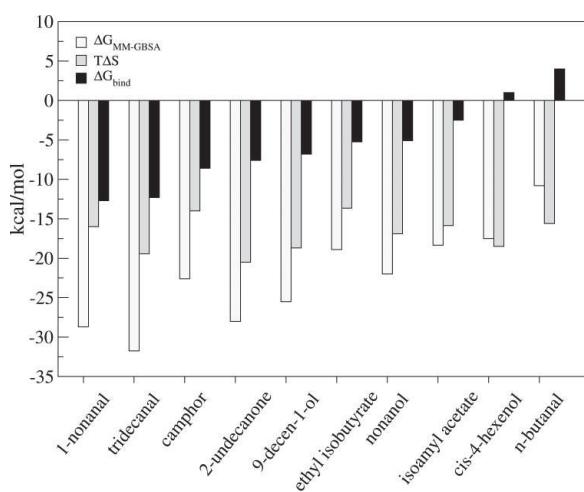


Figure 3. Estimation of the enthalpy ($\Delta G_{MM\text{-}GBSA}$) and entropy ($T\Delta S$) contributions to the free energy of binding (ΔG_{bind}) for 8 agonists and 2 non-agonists of hOR1G1. $\Delta G_{MM\text{-}GBSA}$ values for 1-nonanal, camphor, 9-decen-1-ol, nonanol and n-butanal are taken from ref.[9]

The free energy of binding is split into energy and entropy components. The eight agonists are clearly discriminated from the two non-agonists with an analysis equivalent to that found in Table 1. All agonists free energies are in the range of [-12; -3] kcal·mol⁻¹, which correspond to nanomolar to millimolar affinities. The two non-agonists show positive free energies of binding indicating that they are thermodynamically favored as non-bonded to the receptor in eq. 1. The difference between the weakest agonist and the first non-agonist is 3.5 kcal·mol⁻¹. This illustrates the robustness of the presented protocol to split large sets of odorants into binders and non-binders for a given OR. The discrimination between agonists and antagonists within the binders' family (associated to eq. 2) will require predicting the activation of the receptor. This task is actually in progress.

The combinatorial code of the perception of smell, but more importantly the almost infinite chemical space of odorant molecules, makes it impossible to experimentally identify all combinations of odorants and ORs. Our protocol, based on a thorough sampling of the conformational space of odorant/OR complexes, will help decreasing the number of odorants to be further tested by means of experimental assays. It represents a proof of principle for a robust computational approach which will help unravel odor coding in the nervous system solely based-on the sequence of a receptor and a list of odorants. It will further facilitate the establishment of general rules meant to predict behaviors of the olfactory neurons in our brain, then deciphering the combinatorial code associated to the perception of smell.

References

- [1] T. Bozza, P. Feinstein, C. Zheng, P. Mombaerts, Odorant Receptor Expression Defines Functional Units in the Mouse Olfactory System, *The Journal of Neuroscience* 22 (2002) 3033-3043.
- [2] X. Grosmaitre, A. Vassalli, P. Mombaerts, G.M. Shepherd, M. Ma, Odorant responses of olfactory sensory neurons expressing the odorant receptor MOR23: a patch clamp analysis in gene-targeted mice, *Proceedings of the National Academy of Sciences of the United States of America* 103 (2006) 1970-1975.
- [3] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [4] A. Matsui, Y. Go, Y. Niimura, Degeneration of Olfactory Receptor Gene Repertoires in Primates: No Direct Link to Full Trichromatic Vision, *Molecular Biology and Evolution* 27 (2010) 1192-1200.
- [5] B. Malnic, P.A. Godfrey, L.B. Buck, The human olfactory receptor gene family, *Proceedings of the National Academy of Sciences of the United States of America* 101 (2004) 2584-2589.
- [6] G. Launay, G. Sanz, E. Pajot-Augy, J.F. Gibrat, Modeling of mammalian olfactory receptors and docking of odorants, *Biophysical Reviews* 4 (2012) 255-269.
- [7] U. Gether, B.K. Kobilka, G protein-coupled receptors. II. Mechanism of agonist activation, *The Journal of biological chemistry* 273 (1998) 17979-17982.
- [8] P. Hummel, N. Vaidehi, W.B. Floriano, S.E. Hall, W.A. Goddard, Test of the Binding Threshold Hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93, *Protein Science* 14 (2005) 703-710.
- [9] L. Charlier, J. Topin, C. Ronin, S.K. Kim, W.A. Goddard, 3rd, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cellular and molecular life sciences : CMLS* 69 (2012) 4205-4213.
- [10] V. Matarazzo, C. Ronin, Human olfactory receptors: recombinant expression in the baculovirus/Sf9 insect cell system, functional characterization, and odorant identification, *Methods Mol Biol* 1003 (2013) 109-122.
- [11] G. Sanz, T. Thomas-Danguin, H. Hamdani el, C. Le Poupon, L. Briand, J.C. Pernollet, E. Guichard, A. Tromelin, Relationships between molecular structure and perceived odor quality of ligands for a human olfactory receptor, *Chem Senses* 33 (2008) 639-653.
- [12] G. Sanz, C. Schlegel, J.C. Pernollet, L. Briand, Comparison of odorant specificity of two human olfactory receptors from different phylogenetic classes and evidence for antagonism, *Chem Senses* 30 (2005) 69-80.
- [13] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, Molecular modelling of odorant/olfactory receptor complexes, *Methods Mol Biol* 1003 (2013) 53-65.
- [14] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angewandte Chemie International Edition* 51 (2012) 1274-1278.
- [15] J.H. Park, T. Morizumi, Y. Li, J.E. Hong, E.F. Pai, K.P. Hofmann, H.W. Choe, O.P. Ernst, Opsin, a structural model for olfactory receptors?, *Angew Chem Int Ed Engl* 52 (2013) 11021-11024.
- [16] L. Doszczak, P. Kraft, H.P. Weber, R. Bertermann, A. Triller, H. Hatt, R. Tacke, Prediction of perception: probing the hOR17-4 olfactory receptor model with silicon analogues of bourgeonal and lilial, *Angew Chem Int Ed Engl* 46 (2007) 3367-3371.
- [17] P.C. Lai, B. Guida, J. Shi, C.J. Crasto, Preferential Binding of an Odor Within Olfactory Receptors: A Precursor to Receptor Activation, *Chemical Senses* 39 (2014) 107-123.

- [18] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Science* 13 (2004) 240-254.
- [19] G. Launay, G. Sanz, E. Pajot-Augy, J.-F. Gibrat, Modeling of mammalian olfactory receptors and docking of odorants, *Biophysical Reviews* 4 (2012) 255-269.
- [20] V. Le Guilloux, P. Schmidtke, P. Tuffery, Fpocket: an open source platform for ligand pocket detection, *BMC bioinformatics* 10 (2009) 168.
- [21] P. Schmidtke, V. Le Guilloux, J. Maupetit, P. Tuffery, fpocket: online tools for protein ensemble pocket detection and tracking, *Nucleic acids research* 38 Suppl (2010) W582-589.
- [22] P.C. Lai, C.J. Crasto, Beyond Modeling: All-Atom Olfactory Receptor Model Simulations, *Frontiers in Genetics* 3 (2012)
- [23] A.M. Ring, A. Manglik, A.C. Kruse, M.D. Enos, W.I. Weis, K.C. Garcia, B.K. Kobilka, Adrenaline-activated structure of beta(2)-adrenoceptor stabilized by an engineered nanobody, *Nature* 502 (2013) 575-579.
- [24] R.O. Dror, A.C. Pan, D.H. Arlow, D.W. Borhani, P. Maragakis, Y. Shan, H. Xu, D.E. Shaw, Pathway and mechanism of drug binding to G-protein-coupled receptors, *Proceedings of the National Academy of Sciences* 108 (2011) 13118-13123.

Article 8 - Le spectre de reconnaissance d'un RO est modulé par son affinité avec les odorants mais aussi par sa barrière d'activation.

Le calcul d'affinité odorant/récepteur permet manifestement de faire un premier tri entre les molécules agonistes et non-agonistes d'un RO. Ces récepteurs, qui sont au nombre d'environ 400 chez l'humain, couvrent des spectres de reconnaissance variés. En effet, certains ROs s'activent pour un grand nombre d'odorants et d'autres pour un nombre limité. Ils sont dits à spectre de reconnaissance large ou restreint (Figure 1).

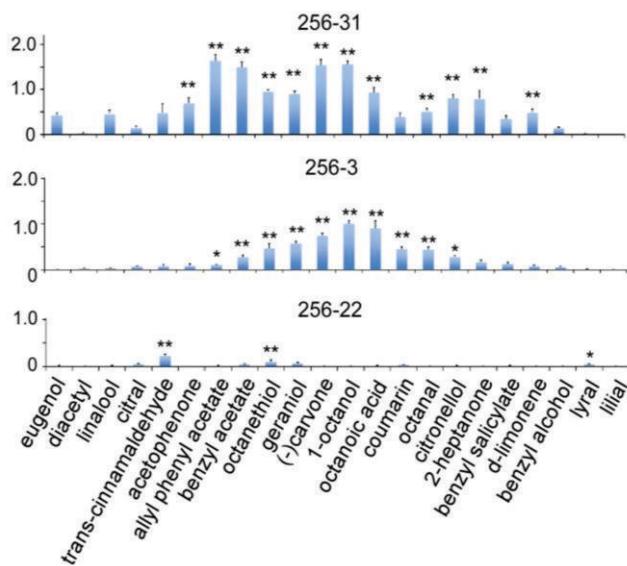


Figure 1. Réponses *in vitro* de différents ROs de la souris (mOR256-31, mOR256-3 et mOR256-22) à un groupe de 22 odorants après soustraction de l'activité basale (activité indépendante d'un ligand). Tous les odorants sont testés à 300 µM. Les données pour chaque ROs sont moyennées sur trois répétitions. Une réponse positive est identifiée par une ou deux étoile(s) si elle est significativement plus élevée que l'activité basale. mOR256-31 et mOR256-3 possèdent des spectres de reconnaissance large, répondant à de nombreux odorants. mOR256-22, quant à lui, ne répond qu'à trois odorants et est donc à spectre de reconnaissance restreint.

Quels sont les mécanismes régulant le spectre de reconnaissance d'un RO ?

Pour cette étude, une série de ROs de la souris est étudiée, celle des mOR256-X. Ils possèdent la particularité de partager une grande identité de séquence mais de posséder des spectres de reconnaissance allant du large au restreint (Figure 1). Cette différence de reconnaissance est donc due à des phénomènes subtils. Il est démontré dans cet article que le spectre de reconnaissance d'un RO est contrôlé à la fois par la permissivité de sa cavité et par sa capacité à s'activer (Figure 2). Ces caractéristiques peuvent être schématisées par un profil d'énergie dépendant de deux

paramètres : 1/ la permissivité de la cavité de liaison de l'OR, 2/ sa capacité à passer d'un état inactif à un état actif (Figure 2).

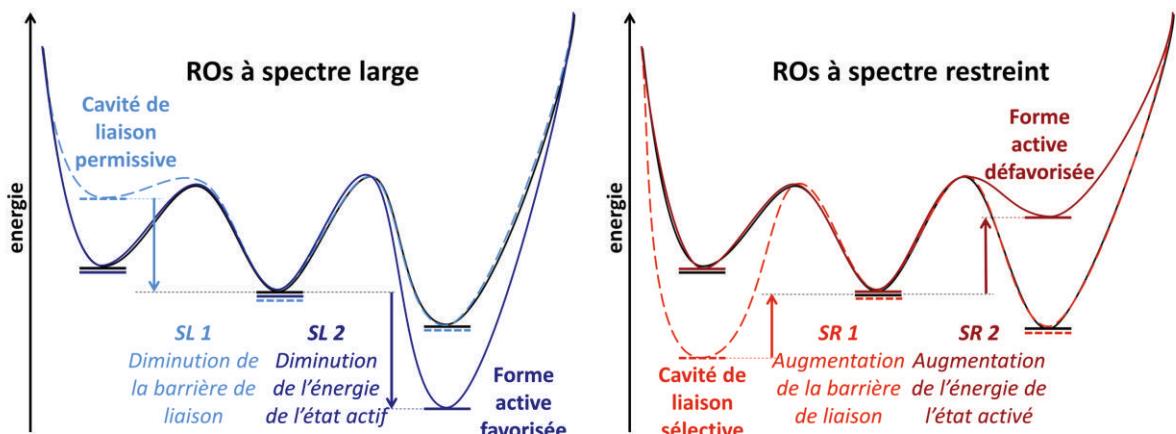


Figure 2. Schéma du profil énergétique du mécanisme d'activation des ROs à spectre large et restreint. Gauche - profils énergétiques correspondant à un RO à large spectre possédant une cavité de liaison permissive (SL1 - bleu clair) ou une faible barrière d'activation (SL2 - bleu foncé). Droite - profils énergétiques correspondant à un RO à spectre restreint possédant une cavité de liaison sélective (SR1 - corail) ou une forte barrière d'activation (SR2 - bordeaux).

Cette hypothèse permet de considérer que l'affinité d'un odorant pour la cavité d'un RO, bien qu'êtant importante (cf Article 6), n'est pas la caractéristique unique pour engendrer l'activation du RO et donc une réponse neuronale.

Dans cet article, j'ai réalisé toutes les expériences et analyses de modélisation moléculaire : construction des modèles, réalisation des simulations de dynamique moléculaire, calcul de l'enthalpie libre de liaison et décomposition par résidus de cette énergie. Yiqun Yu (University of Pennsylvania) était en charge de la partie expression des ROs, réalisation des mutants et mesure des réponses des ROs par des odorants. Nous avons travaillé en collaboration pour le choix des résidus à muter pour l'étude.

Article 8:

Broad Responsiveness of G Protein-Coupled Odorant Receptors Is Attributed to the Activation Mechanism

Yiqun Yu[#], Claire A de March[#], Mengjue J Ni, Kaylin A. Adipietro, Jérôme Golebiowski, Hiroaki Matsunami, Minghong Ma, submitted to PNAS (2015)

These authors contributed equally to this work.

Keywords: G protein-coupled receptor; odorant receptor; broad response profile; site-directed mutagenesis; computational modeling

Abstract

Mammals detect and discriminate numerous odors via a large family of G protein-coupled odorant receptors (ORs). However, little is known about the molecular and structural basis underlying OR response properties. Using site-directed mutagenesis and computational modeling, we studied ORs sharing high sequence homology but with different response properties. When tested in heterologous cells by diverse odorants, MOR256-3 responded broadly to many odorants while MOR256-8 responded weakly to a few odorants. Out of 36 mutant MOR256-3 ORs, the majority altered the responses to different odorants in a similar manner and the overall response of an OR is positively correlated with its basal activity, an indication of ligand-independent receptor activation. Strikingly, a single mutation in MOR256-8 is sufficient to confer both high basal activity and broad responsiveness to this receptor. These results suggest that broad responsiveness of an OR is at least partially attributed to its activation mechanism.

Introduction

G protein-coupled receptors (GPCRs) are seven transmembrane (TM) proteins which play essential roles in converting extracellular stimuli into intracellular signals in a variety of cell types. Odor detection by olfactory sensory neurons (OSNs) in the mammalian nose depends on a large family of G protein-coupled odorant receptors (ORs) [1], which endows the olfactory system with an extraordinary power of odor detection and discrimination. Although OR-ligand binding is the first step towards smell perception, little is known about the molecular and structural basis underlying odor response properties of individual ORs.

Most mammalian ORs respond to a small fraction of all the tested odorants [2]. In contrast, recent studies have identified a small number of ORs that respond to a large set of diverse odorants with comparable potency and efficacy as the former. Curiously, several broadly responsive ORs including MOR256-3 (Olfr124 or SR1), MOR256-31 (Olfr263), and human OR2W1 (ortholog of MOR256-31) belong to the same subfamily, which also contains ORs such as MOR256-8 (Olfr1362) and MOR256-22 (Olfr1387) that respond to a few odorants [3-6]. Identification of ORs within the same subfamily (i.e., sharing >50% amino acid identity) but with different response properties offers an opportunity for dissecting out the molecular features that define the tuning properties of these ORs.

Mammalian ORs belong to class A (or *rhodopsin* family) GPCRs. The structure-function relationship of several class A members (e.g., rhodopsin and β 2-adrenergic receptor) has been investigated in great details via various approaches including site-directed mutagenesis, X-ray crystallography [7, 8], and molecular modeling [9-11]. Although no crystal structure is available for any OR, site-directed mutagenesis and/or computational modeling have shed light on structure-function relationship for a

few ORs [12-17]. Our recent study reveals the critical role of a few conserved residues in either ligand binding or receptor activation using MOR256-3 as a model [18].

Using a joint approach of site-directed mutagenesis and computational modeling, we investigated the response properties of mutant ORs based on MOR256-3 and MOR256-8, which respond to a large and small set of odorants, respectively. Three-dimensional atomic models of these ORs were built to map locations of the mutated residues. Most mutations in MOR256-3 altered the responses to different odorants in a similar manner. Remarkably, MOR256-8 was converted into a broadly responsive OR by swapping a single or a few residues. More generally, we found that an OR's total response is positively correlated with its basal activity, an indication of ligand independent receptor activation. These data suggest that broad responsiveness of an OR is not only determined by ligand binding, but also by activation mechanism.

Results

Identification of key residues that potentially underlie broad responsiveness of MOR256-3

Ideally the response profile of an OR should be determined based on an exhaustive list of odorants, which would be time consuming if not impossible given the almost infinite odor space. To provide a numerical description of broadly responsive ORs, we analyzed the percentage of odorants a receptor responds to from an array of 62 ORs from different subfamilies vs 63 diverse odorants reported in a previous study [5]. The median is 4.8% and the median absolute deviation (MAD) is 7.0%, indicating that any OR responding to 25.8% of the odorants would be 3 MAD away from the median. We hence define an OR as broadly responsive if it responds to >30% of a given set of diverse odorants, which covers a significant portion of the odor space [19]. This definition offers an appropriate description of ORs with exceptionally broad response profiles using different sets of odorants (Fig. S1). To minimize the effects of odorant concentrations on the OR response profiles, all ORs were tested at the same concentrations and a positive response was determined at 300 μ M, a near-saturating concentration.

We initially focus on the broadly responsive MOR256-3 receptor which has been extensively studied both in genetically-tagged OSNs and in a heterologous expression system [3]. We compared the response properties of the following five ORs within the same family. Out of 22 diverse odorants tested, MOR256-3, MOR256-31 and hOR2W1 exhibited broad responsiveness. In contrast, MOR256-8 and MOR256-22 responded weakly to a few odorants (Fig. 1A). Their response profiles were further assessed by dose-response analysis on selected odorants or a larger odorant set (Fig. S1 and [5]). Note that odorant-induced responses are not correlated with the receptor surface expression levels (see below).

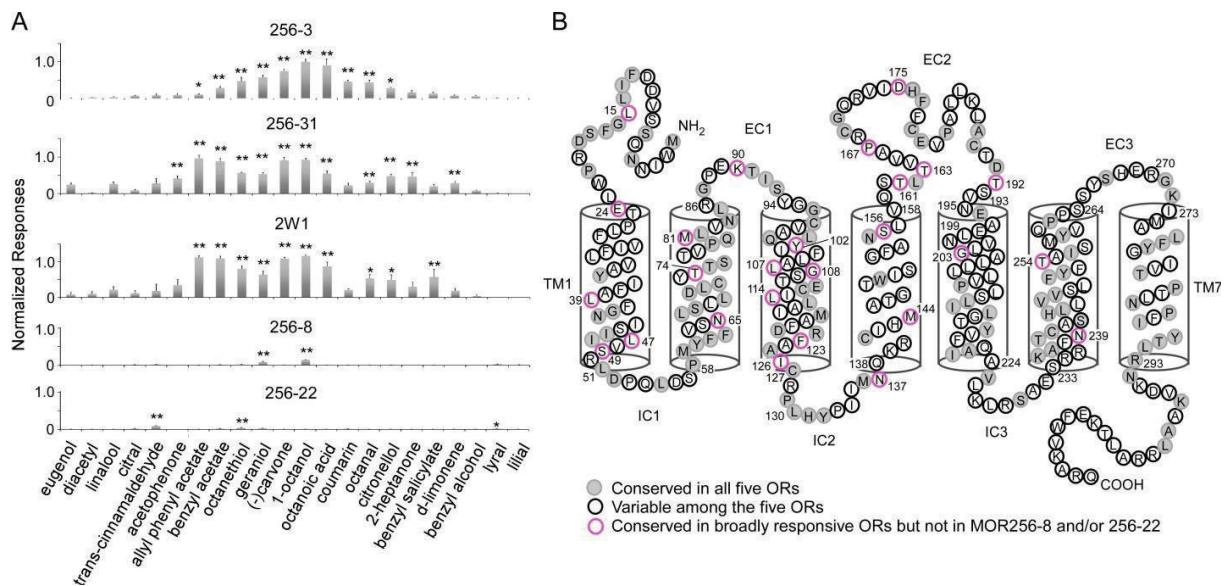


Figure 1. The MOR256 subfamily contains ORs with different response properties. (A) Responses of different ORs to a set of 22 odorants (all at 300 μ M) in Hana3A cells (mean \pm s.e.m.). All odorant-OR pairs were tested on at least two plates (with three repeats on each plate). A positive response is identified if it is significantly higher than the basal activity (* $p<0.05$ and ** $p<0.01$ in one-way ANOVA post hoc tests). MOR256-3, MOR256-31 and 2W1 responded to 10, 12 and 10 compounds (or 45.5%, 54.5%, and 45.5%), respectively. MOR256-8 and MOR256-22 showed weak, but significant responses to two and three odorants, respectively, due to their low basal activity. All responses were normalized to MOR256-3's response to 1-octanol at 300 μ M and corrected for surface expression (see *Materials and Methods* for details). (B) Snake plot of the MOR256-3 receptor, which contains 315 amino acids with 114 conserved in all five ORs (filled in gray). The transmembrane domains are determined by the 3D atomic model (see also Fig. 2). Magenta circles mark residues that are conserved in the three broadly responsive ORs (MOR256-3, MOR256-31 and 2W1), but not in MOR256-8 and/or MOR256-22.

In order to identify key residues that underlie broad responsiveness, we aligned the protein sequences of these five ORs from the same subfamily (Fig. 1B and Fig. S2) and built 3D atomic models of MOR256-3 and MOR256-8 (Fig. 2B) [18, 20]. We constructed 36 site-directed mutant MOR256-3 ORs mostly by substituting the residues conserved in the broadly responsive ORs to those in MOR256-8 or MOR256-22 individually or in combination. The mutated sites included all 17 conserved residues between TM3 and TM6 plus six located in TM1 and TM2 (Fig. 1B).

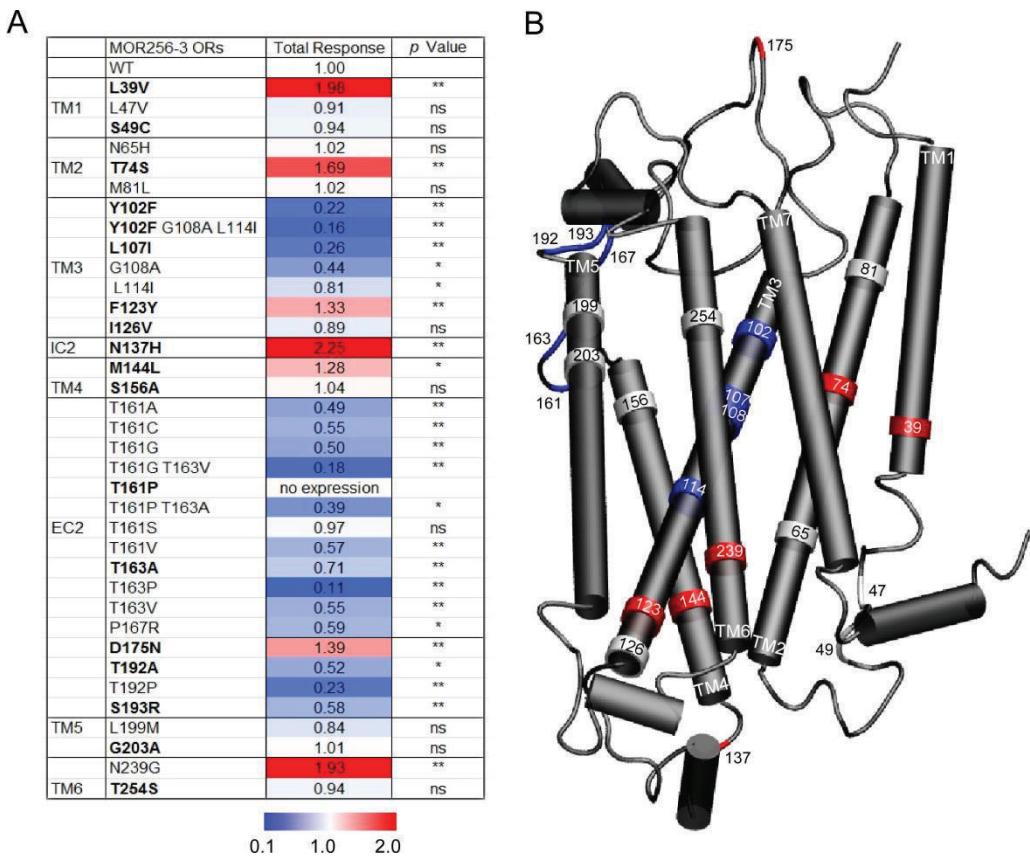


Figure 2. Summary of the total response of mutant MOR256-3 ORs. (A) The total response is the sum of the responses to all five odorants at 300 μ M, normalized to that of WT MOR256-3 tested on the same plate and corrected for surface expression. TM = transmembrane domain; IC = intracellular loop; and EC = extracellular loop. The mutants in bold mark residue swaps between MOR256-3 and MOR256-8. One-way ANOVA post hoc tests were performed for each mutant and WT pair (ns = not significantly different, * $p < 0.05$, and ** $p < 0.01$). The responses to individual odorants are reported in Fig. 4A and Supplemental Table S1. (B) Mutated residues are shown in the 3D atomic model of MOR256-3. Red, blue and gray colors indicate increased, decreased, and unchanged total response, respectively. The MOR256-8 model looks almost identical to the MOR256-3 model.

All MOR256-3 mutants except T161P are expressed at the cell surface (Fig. S3) and differentially influence odorant-induced responses (Figs. 2, 3 and Table S1). Eighteen mutations significantly decreased, seven significantly increased, and ten did not change the overall responses (Figs. 2A, B, 3). Notably, switching a single residue in TM3 of MOR256-3 to that of MOR256-8 (denoted as 3 Y102F or 3 L107I) drastically decreased the odor responses by >70% (Figs. 2A, 3, 4A). When the responses to individual odorants at 300 μ M were ranked, most mutant ORs showed the same ranking order as wild-type (WT) 256-3 (from the strongest to weakest ligand: 1-octanol, (-) carvone, coumarin, benzyl acetate, and allyl phenyl acetate) with a few exceptions (G108A, L199M, G203A, and T254S) (Figs. 3, 4A), suggesting that most of the mutated residues are not governing binding to specific odorants but rather affecting the overall responsiveness.

Curiously, ORs with strong odorant responses statistically showed higher basal activities (Fig. 4A). Regression analysis on the data set including the five WT ORs and all 35 functional mutant MOR256-3 ORs confirmed that the total response of an OR is positively correlated with the basal activity (Fig. 4B). In contrast, neither the total response nor the basal activity is correlated with the OR surface expression level (Fig. 4C). These data support that more responsive ORs have a higher basal activity level, implying a higher probability of receptor activation.

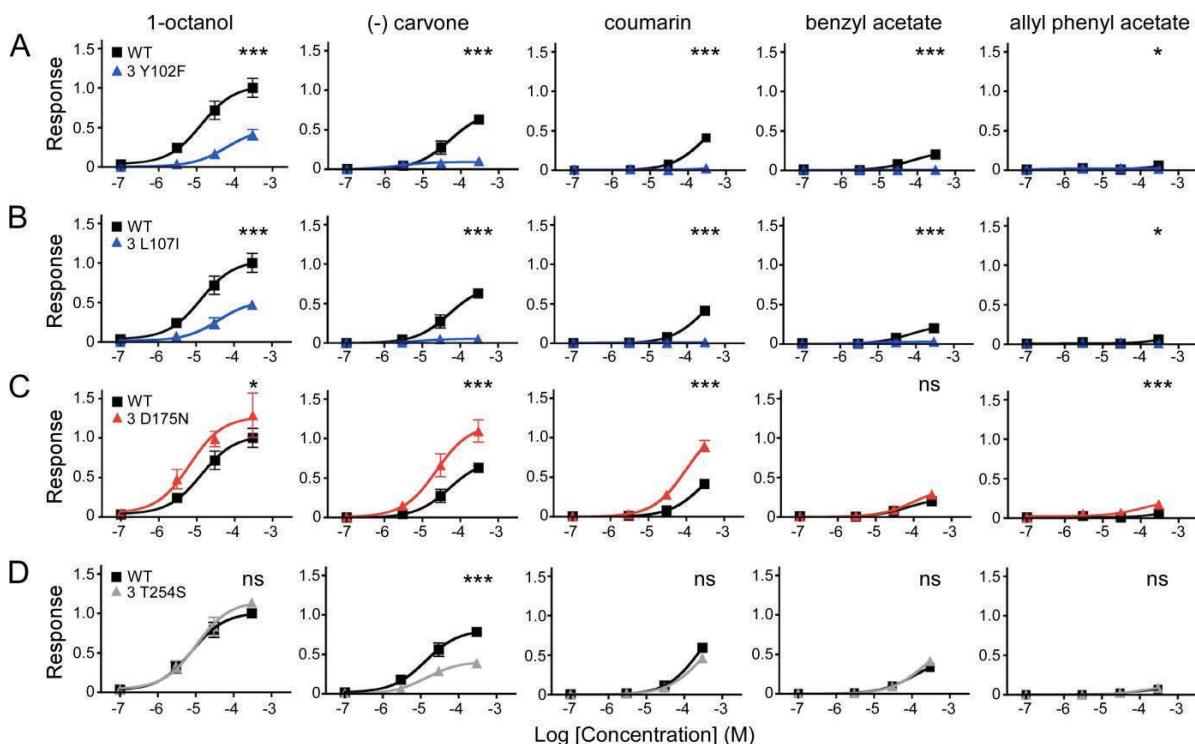


Figure 3. Mutations at different residues differentially change the response profiles of MOR256-3. (A, B) Single mutations 3 Y102F (A) and 3 L107I (B) decreased the responses to all five odorants. (C) A single mutation 3 D175N increased the responses to four out of five odorants. (D) A single mutation 3 T254S selectively reduced the response to (-) carvone. Each mutant OR was tested on the same plate as WT (three repeats for each OR) and all responses were normalized to WT response to 1-octanol at 300 μ M and corrected for surface expression. Two-way ANOVA (concentration and OR type) tests were performed for each mutant and WT pair (ns = not significantly different, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$ for OR type).

A single mutation confers broad responsiveness to MOR256-8

MOR256-8 shares more than 50% amino acid identity with other broadly responsive members in the same subfamily. Because in MOR256-3, substituting single residues by those in MOR256-8 can lead to complete loss of surface expression (3 T161P) or significantly reduced odorant responses (3 Y102F and 3 L107I) (Figs. 2-4, S3), we asked whether reversely swapping these residues would confer broad responsiveness to MOR256-8. All mutant MOR256-8 ORs described below showed surface expression and their responses are not correlated with the expression levels (Fig. 5C).

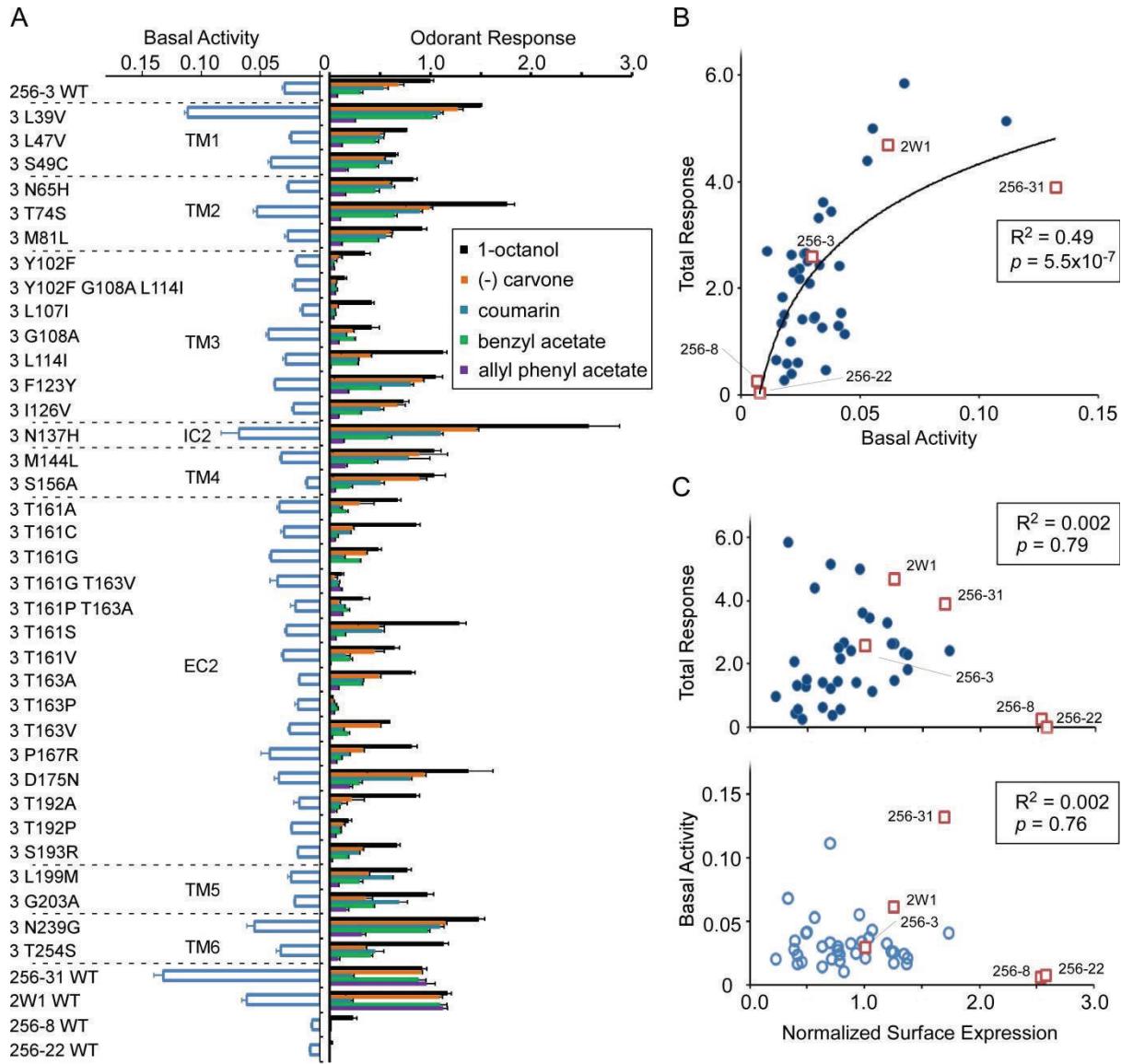


Figure 4. The total response of an OR is positively correlated with its basal activity.

(A) Summary of the basal activity (left) and the responses to all five odorants at 300 μ M (right) for each OR. The five odorants are ranked based on the responses of the WT MOR256-3 from largest to smallest. The odorant responses for each OR were averaged from three repeats on the same plate except for WT ORs, averaged from 8-21 plates (mean \pm s.e.m.). The basal activity for each OR was averaged from four repeats on the same plate. All odorant responses and basal activities were normalized to WT MOR256-3's response to 1-octanol at 300 μ M and corrected for surface expression. (B) The total response to all five odorants at 300 μ M is plotted against its basal activity for each OR. The curved line represents logarithmic regression fitting because it models the fact that the odorant-induced response cannot rise linearly but instead reaches a plateau even for an extremely broadly responsive OR. (C) Neither the total response nor the basal activity of WT and mutant ORs is correlated with the receptor surface expression (see Figure S3 for the surface expression of each OR) via linear regression analysis.

Single mutations 8 F102Y and 8 P161T responded to three and six odorants, respectively, more than WT MOR256-8, which responded to two out of the 22 odorants. Strikingly, mutations 8 I107L and 8 I107L P161T responded to 45.5% and 40.9% of the odorant set, respectively, indicating that they are broadly responsive (Fig. 5A). Triple mutation 8 F102Y I107L P161T did not respond more broadly than 8 I107 and 8 I107L P161T (Fig. 5A), indicating that the effects of these residues are not additive. Compared with WT MOR256-8, the broadly responsive mutant ORs showed higher basal activity (Fig. 5B) as other broadly responsive ORs (Fig. 4).

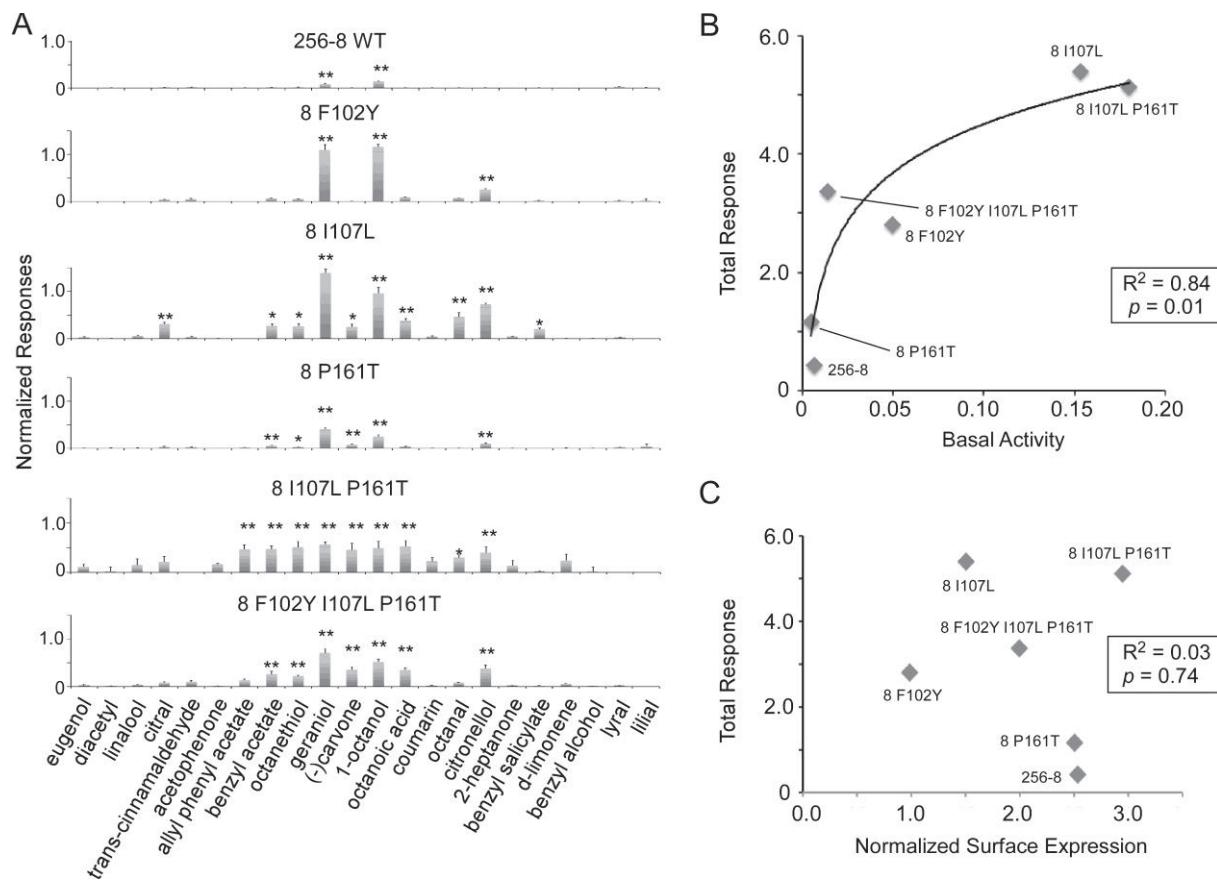


Figure 5. Mutations at a single or a few residues are sufficient to convert the narrowly-tuned MOR256-8 to broadly-tuned ORs. (A) Responses of different ORs to a set of 22 odorants after subtracting the basal activity. All odorants were at 300 μ M and the data for each OR were averaged from three repeats on the same plate (mean \pm s.e.m.). All odorant responses and basal activities were normalized to WT MOR256-3's response to 1-octanol at 300 μ M and corrected for surface expression. A positive response is identified if it is significantly higher than the basal activity (* $p < 0.05$ and ** $p < 0.01$ in one-way ANOVA post hoc tests). (B) The total response (sum of the responses to all 22 odorants at 300 μ M) is plotted against its basal activity for each OR. All activities were normalized to WT MOR256-3's response to 1-octanol at 300 μ M. The curved line represents logarithmic regression fitting. (C) The total response is not correlated with the receptor surface expression (linear regression analysis). The surface expression of each OR is normalized to that of WT MOR256-3.

Discussion

In the current study, we investigated the molecular and structural features underlying the broad responsiveness of MOR256-3 using site-directed mutagenesis and computational modeling. We identified a handful of residues in MOR256-3 that are critical for its response properties (Figs. 1-3). We further demonstrate that the basal activity of an OR is strongly correlated with its overall responsiveness (Fig. 4). Remarkably, MOR256-8, which weakly responds to a few odorants, can be converted into a broadly responsive receptor by mutation of a single or a few residues (Fig. 5). These data provide new insights into the mechanisms underlying the response properties of mammalian ORs and GPCRs in general.

Our study suggests that the activation mechanism of an OR significantly impacts its response properties. The major difference between MOR256-3 and MOR256-8 likely resides in the receptor activation process. Many single mutations in MOR256-3 changed its response efficacy and potency to all odorants in a similar manner (Figs. 2-4). Strikingly, mutation of a single (at position 107) or a few residues in MOR256-8 confers both high basal activity and broad responsiveness to this receptor. Interestingly, residue 107 is at the vicinity of G108 (identical between MOR256-3 and MOR256-8) and mutation G108L in MOR256-3 leads to a constitutively active receptor [18]. This position is located at the crossing of TM3 and TM6 and is reported to form a cradle for the ligand in class A GPCRs [21]. It is plausible that mutation at 107 and many other residues examined in this study change the likelihood of receptor activation rather than binding to specific ligands.

A broadly responsive OR (such as MOR256-3) would require a low activation barrier in addition to a permissive binding pocket. Unlike most non-olfactory GPCRs which utilize ionic or hydrogen bond interactions to ensure selectivity and sensitivity in binding their agonists or antagonists [22], ORs bind to their ligands via much weaker van der Waals interactions [14, 16, 23]. It is conceivable that at least some ORs have a permissive binding cavity where the interaction would be more opportunistic compared to other GPCRs [14]. Therefore, binding of many distinct ligands can lead to receptor activation.

On the other hand, an OR that appears as “selectively responsive” could result from two scenarios: either this OR has a truly restrictive binding cavity or it has a high activation barrier for activation. We suspect that MOR256-8 belongs to the latter since mutation of a single or a few residues can convert it into a broadly responsive OR (Fig. 5).

The positive correlation between broad responsiveness and basal activity described here (Figs. 4, 5) does not imply that the basal activity alone is sufficient to predict the response properties of a given OR. This correlation may only apply to some ORs and operate within a certain range. For example, a constitutively active mutant OR (with extremely high basal activity) can lose its capability of

responding to odorants [18]. Many residues in addition to the ones investigated here may also contribute to an OR's response properties by affecting ligand binding and/or receptor activation.

Broadly responsive ORs are identified from insects to humans [5, 24], and the selective advantages of these ORs in smell perception may be multifaceted. First, these receptors can potentially increase the range of detectable odors for the olfactory system. Second, they may contribute to an organism's ability in odor discrimination via the combinatorial scheme. In the retina, three types of cones with broad tuning spectra are sufficient for color perception. Third, these receptors may serve as general odor detectors. Curiously, OSNs expressing MOR256-3 are highly concentrated in the septal organ, a chemosensory organ located in the ventral base of the nasal septum and in the direct air path [25]. The broadly responsive OSNs in the septal organ and the main olfactory epithelium may inform the system the presence of any odor in the environment. Fourth, these receptors may act as intensity analyzers by providing the olfactory system an easier readout on odor concentrations regardless of the identity of the odors.

Materials and Methods

Site-Directed Mutagenesis

The coding sequences of MOR256 receptors were amplified from genomic DNA of C57BL/6 mice and subcloned into the pcDNA3.1/TOPO vector (Invitrogen) with an N-terminal tag of the first 20 amino acids of rhodopsin (Rho). Site-directed mutants were constructed using the Quikchange site-directed mutagenesis kit (Agilent Technologies). The sequences of all plasmid constructions were verified by both forward and reverse sequencing (DNA sequencing core facility, University of Pennsylvania).

Luciferase assay in Hana3A cells

The Dual-Glo Luciferase Assay (Promega) was used to determine the activities of firefly and Renilla luciferase in Hana3A cells [26]. Firefly luciferase, driven by a cAMP response element promoter (CRE-Luc; Stratagene), was used to determine OR activation levels. Renilla luciferase, driven by a constitutively active SV40 promoter (pRLSV40; Promega), functioned as an internal control for transfection efficiency and cell viability. Hana3A cells, a HEK293T-derived cell line stably expressing the receptor-transporting proteins (RTP1L and RTP2), receptor expression-enhancing protein 1 (REEP1), and olfactory G protein (G_{\alphaolf}) [26], were plated on poly-D-lysine-coated 96-well plates (Nalge Nunc) and incubated overnight in minimum essential medium eagle (Sigma) with 10% FBS at 37°C and 5%CO₂. The following day, cells were transfected using Lipofectamine 2000 (Invitrogen). For each 96-well plate, 0.5 µg pRL-SV40, 1 µg CRE-Luc, 0.5 µg mouse RTP1S, 0.25 µg mouse muscarinic acetylcholine receptor M3 and 0.5 µg of receptor plasmid DNA were transfected. After transfection (24 h), medium was replaced with 25 µl of odorant solution diluted in CD293 chemically defined

Article 8 - Yu, de March et al. soumis à PNAS (2015)

medium (Invitrogen), and cells were further incubated for 4 h at 37°C and 5% CO₂. The manufacturer's protocols were followed to measure firefly luciferase and Renilla luciferase activities. Raw data were analyzed according to the published procedure using Microsoft Excel and GraphPad Prism [26].

To facilitate comparison between OR responses from multiple plates, we always included Rho-tag empty vector and WT MOR256-3 as negative and positive control, respectively. The basal activity of an OR was averaged from four wells in the absence of odorants. An odorant-induced response was obtained by subtracting the basal activity of that receptor. In experiments where the five odorants (1-octanol, (-) carvone, coumarin, benzyl acetate, and allyl phenyl acetate) were tested, the responses to each odorant at four concentrations (0, 3, 30, and 300 μM) were measured. The sum of the responses to all five odorants at 300 μM was used to evaluate an OR's overall responsiveness. All odorant responses and basal activities were normalized to MOR256-3's response to 300 μM 1-octanol.

Evaluation of OR surface expression and data correction

Live-cell immuno-staining was used to evaluate OR surface expression [26]. Hana3A cells were co-transfected with receptor and GFP plasmids 24 hours before the staining. The transfected Hana3A cells were incubated with primary antibody solution (mouse anti-rhodopsin, Rho 4D2, Abcam) on ice for 1 h. After rinsing the cells for three times, secondary antibody solution (Alexa Fluor 568-conjugated anti-mouse IgG) was added onto the cells, and incubated for 45 min on ice. At the end of the incubation, the cells were fixed with 2% paraformaldehyde, and mounted with vectashield mounting medium (Vector Laboratories, Inc.).

The surface expression of each OR was quantified by the Rho/GFP intensity ratio, which takes into account both the number of transfected cells and the expression level in individual cells. For each plate, the total fluorescence intensity (after background subtraction) was measured for both red (Rho) and green (GFP) channel using Image J. The intensity ratio (Rho/GFP) was obtained for all WT and mutant ORs (Fig. S3). All odorant response and basal activities were corrected for OR surface expression by dividing the Rho/GFP intensity ratio scaled to WT MOR256-3.

For selected ORs (Fig. S3), the surface expression was also evaluated using fluorescence activated cell sorting (FACS). The transfected cell pellets were resuspended in primary antibody solution (mouse anti-Rhodopsin), and incubated on ice for 1 h. After centrifuging for 3 min at 200g and aspirating all solution, the cell pellets were resuspended in secondary antibody solution (Phycoerythrin-conjugated anti-mouse IgG), and incubated for 30 min on ice. At the end of the incubation, the cells were centrifuged, and the cell pellets were suspended with washing solution in 5-ml round-bottomed tubes (BD Falcon). Cells were analyzed using a FACS machine (Flow Cytometry and Cell Sorting

Resource Laboratory, University of Pennsylvania) according to the GFP and PE fluorescent signals. The fluorescent range of the GFP and PE were determined by control cells transfected with GFP or receptor plasmid only. Cells transfected with GFP only were also used to determine the nonspecific PE fluorescence.

3D atomic models

The protocol follows a previously published method [27]. Sequences of MOR256-3, 256-8, 256-17, 256-22, 256-31, mI7 (olfr2), mOR-EG (olfr73), and S25 (olfr480) are aligned with 396 human ORs [28] and nine sequences of X-ray elucidated GPCRs: bovine rhodopsin (Protein Data Bank or PDB: 1U19) [29], human beta 2 adrenergic (PDB: 2RH1) [30], turkey beta 1 adrenergic (PDB: 2VT4) [31], human chemokine receptors CXCR4 (PDB: 3ODU) [32] and CXCR1 (PDB: 2LNL) [33], human dopamine receptor D3 (PDB: 3PBL) [34], human adenosine a2A receptor (PDB: 2YDV) [35], human histamine H1 receptor (PDB: 3RZE) [36] and muscarinic acetylcholine receptor M2 (PDB: 3UON) [37]. Highly conserved motifs in ORs are considered as constraints for the alignment: GN in helix 1, PMYFFLXXLSXXD in helix 2, MAYDRYXAICXPLXY in helix 3, SYXXI in helix 5, KAFSTCASH in helix 6, LNPXIY in helix 7 and a pair of conserved cysteines 97^{3.25}-179^{EC2} which constitute a known disulfide bridge between the beginning of helix 3 and the extracellular loop 2. Four experimental GPCR structures (1U19, 3ODU, 2YDV and 2LNL) are selected as templates to build MOR256-3 by homology modeling with Modeller [38]. The N-terminal structure (residues 1 to 18) is excluded to avoid perturbation of the modeling protocol with a non-structured part of the protein.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: a molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] K. Touhara, L.B. Vosshall, Sensing odorants and pheromones with chemosensory receptors, *Annu Rev Neurosci* 71 (2009) 307-332.
- [3] X. Grosmaire, S.H. Fuss, A.C. Lee, K.A. Adipietro, H. Matsunami, P. Mombaerts, M. Ma, SR1, a mouse odorant receptor with an unusually broad response profile, *J Neurosci* 29 (2009) 14545-14552.
- [4] K. Nara, L.R. Saraiva, X. Ye, L.B. Buck, A large-scale analysis of odor coding in the olfactory epithelium, *J Neurosci* 31 (2011) 9179-9191.
- [5] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, J.D. Mainland, Odor coding by a Mammalian receptor repertoire, *Sci Signal* 2 (2009) ra9.
- [6] J. Li, R. Haddad, S. Chen, V. Santos, C.W. Luetje, A broadly tuned mouse odorant receptor that detects nitrotoluenes, *J Neurochem* 121 (2012) 881-890.
- [7] B. Kobilka, The structural basis of G-protein-coupled receptor signaling (Nobel Lecture), *Angewandte Chemie* 52 (2013) 6380-6388.
- [8] R.J. Lefkowitz, A brief history of G-protein coupled receptors (Nobel Lecture), *Angewandte Chemie* 52 (2013) 6366-6378.
- [9] R.O. Dror, D.H. Arlow, P. Maragakis, T.J. Mildorf, A.C. Pan, H. Xu, D.W. Borhani, D.E. Shaw, Activation mechanism of the beta2-adrenergic receptor, *Proc Natl Acad Sci U S A* 108 (2011) 18684-18689.
- [10] N. Vaidehi, W.B. Floriano, R. Trabanino, S.E. Hall, P. Freddolino, E.J. Choi, G. Zamanakos, W.A. Goddard, 3rd, Prediction of structure and function of G protein-coupled receptors, *Proc Natl Acad Sci U S A* 99 (2002) 12622-12627.
- [11] K.J. Kohlhoff, D. Shukla, M. Lawrenz, G.R. Bowman, D.E. Konerding, D. Belov, R.B. Altman, V.S. Pande, Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways, *Nat Chem* 6 (2014) 15-21.
- [12] T. Abaffy, A. Malhotra, C.W. Luetje, The molecular basis for ligand specificity in a mouse olfactory receptor: a network of functionally important residues, *The Journal of biological chemistry* 282 (2007) 1216-1224.
- [13] P.C. Lai, M.S. Singer, C.J. Crasto, Structural activation pathways from dynamic olfactory receptor-odorant interactions, *Chemical senses* 30 (2005) 781-792.
- [14] L. Charlier, J. Topin, C. Ronin, S.K. Kim, W.A. Goddard, 3rd, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cell Mol Life Sci* 69 (2012) 4205-4213.
- [15] C.A. de March, J. Golebiowski, A computational microscope focused on the sense of smell, *Biochimie* 107 Pt A (2014) 3-10.
- [16] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural basis for a broad but selective ligand spectrum of a mouse olfactory receptor: mapping the odorant-binding site, *J Neurosci* 25 (2005) 1806-1815.
- [17] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a ligand-binding niche within a human olfactory receptor by combining site-directed mutagenesis with dynamic homology modeling, *Angewandte Chemie* 51 (2012) 1274-1278.
- [18] C.A. de March, Y. Yu, M.J. Ni, K.A. Adipietro, H. Matsunami, M. Ma, J. Golebiowski, Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors, *J Am Chem Soc* 137 (2015) 8611-8616.

- [19] R. Haddad, R. Khan, Y.K. Takahashi, K. Mori, D. Harel, N. Sobel, A metric for odorant comparison, *Nature methods* 5 (2008) 425-429.
- [20] C.A. de March, S.K. Kim, S. Antonczak, W.A. Goddard, 3rd, J. Golebiowski, G protein-coupled odorant receptors: From sequence to structure, *Protein Sci*, 10.1002/pro.2717(2015)
- [21] A.J. Venkatakrishnan, X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, M.M. Babu, Molecular signatures of G-protein-coupled receptors, *Nature* 494 (2013) 185-194.
- [22] W.A. Goddard, 3rd, R. Abrol, 3-Dimensional structures of G protein-coupled receptors and binding sites of agonists and antagonists, *J Nutr* 137 (2007) 1528S-1538S.
- [23] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The mouse eugenol odorant receptor: structural and functional plasticity of a broadly tuned odorant binding pocket, *Biochemistry* 50 (2011) 843-853.
- [24] E.A. Hallem, J.R. Carlson, Coding of odors by a receptor repertoire, *Cell* 125 (2006) 143-160.
- [25] H. Tian, M. Ma, Molecular organization of the olfactory septal organ, *J Neurosci* 24 (2004) 8383-8390.
- [26] H. Zhuang, H. Matsunami, Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells, *Nature protocols* 3 (2008) 1402-1413.
- [27] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, Molecular modelling of odorant/olfactory receptor complexes, *Methods Mol Biol* 1003 (2013) 53-65.
- [28] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, *Genome Biol* 2 (2001) RESEARCH0018.
- [29] T. Okada, M. Sugihara, A.N. Bondar, M. Elstner, P. Entel, V. Buss, The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure, *J Mol Biol* 342 (2004) 571-583.
- [30] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka *et al*, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science* 318 (2007) 1258-1265.
- [31] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, G.F. Schertler, Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature* 454 (2008) 486-491.
- [32] B. Wu, E.Y. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi *et al*, Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, *Science* 330 (2010) 1066-1071.
- [33] S.H. Park, B.B. Das, F. Casagrande, Y. Tian, H.J. Nothnagel, M. Chu, H. Kiefer, K. Maier, A.A. De Angelis, F.M. Marassi *et al*, Structure of the chemokine receptor CXCR1 in phospholipid bilayers, *Nature* 491 (2012) 779-783.
- [34] E.Y. Chien, W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov *et al*, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, *Science* 330 (2010) 1091-1095.
- [35] G. Lebon, T. Warne, P.C. Edwards, K. Bennett, C.J. Langmead, A.G. Leslie, C.G. Tate, Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation, *Nature* 474 (2011) 521-525.
- [36] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han *et al*, Structure of the human histamine H1 receptor complex with doxepin, *Nature* 475 (2011) 65-70.

PARTIE 3 : Relations structure-fonction des récepteurs olfactifs

Article 8 - Yu, de March et al. soumis à PNAS (2015)

[37] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga *et al*, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, *Nature* 482 (2012) 547-551.

[38] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using Modeller, *Curr Protoc Bioinformatics Chapter 5* (2006) Unit 5 6.

Supporting Information

A

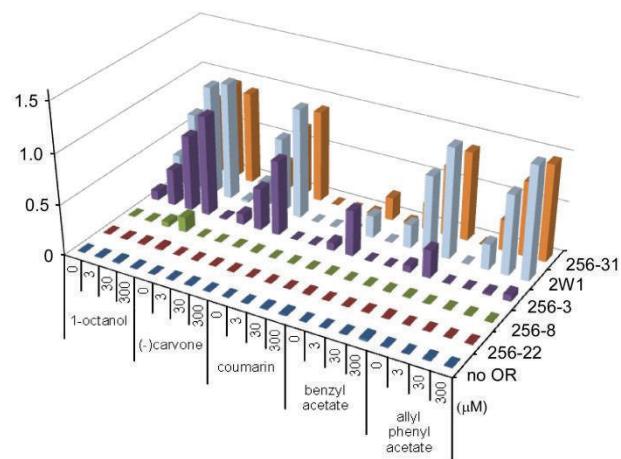


Figure S1. The MOR256 subfamily contains ORs with different response properties. (A) Wild-type MOR256 ORs exhibited different sensitivity and selectivity to the five selected odorants. The data for each receptor were averaged from 8-21 plates (error bars not shown). All responses were normalized to MOR256-3's response to 1-octanol at 300 μ M and corrected for surface expression. (B) Screening of 83 compounds in Hana3A cells revealed 35 ligands (42.2%) for MOR256-3. A positive response was defined when an odorant-induced response was significantly higher than that of the control PCI vector or the basal activity ($p < 0.05$ in one way ANOVA post hoc tests). All compounds were tested at 100 or 300 μ M with at least three repeats.

B

Ligands for MOR256-3	Non-effective Compounds
(-) Carvone	(R)-(+) -Pulegone
1-Octen-3-one	1-Butanol
2,3-Butanediol	2,3-Hexanedione
Allyl Phenyl Acetate	2,5-Dimethyl pyrazine
Amyl propionate	2,5-Dimethylpyrrole
Benzaldehyde	2-heptanone
Benzophenone	2-Methoxy-4-methylphenol
Benzyl acetate	2-Octanone
Citronella	3-Methyl-2-buten-1-ol
Coumarin	Acetaldehyde
E-2-hexen-1-al	Acetophenone
Ethyl octanoate	Allyl hexanoate
Geraniol	Ambrette
Guaiacol	Ammonium hydroxide
Heptanal	Amyl butyrate
Heptanoic acid	Amyl laurate
Heptanol	α -Phellandren
Hexanol	Benzyl alcohol
Hexyl acetate	Benzyl salicylate
Hexyl butyrate	β -Phenethylamine
Hexylamine	Citral
Isoamyl acetate	Cyclohexylamine
Lime	Diacetyl
Methyl eugenol	Diethyl sebacate
n-Hexane	d-Limonene
Nonanal	Ethanol
Nonanoic acid	Ethyl acetate
Octanal	Eugenol
Octanethiol	Furfural
Octanoic acid	Hexyl octanoate
Octanol	Isobutylamine
p-Cymene	Isobutyraldehyde
Piperidine	Isobutyric acid
Undecanoic gamma-lactone	Lilial
	Linalool
	Lyral
	m-Cresol
	Methyl piperidine
	Musk ketone
	Pentadecalactone
	Propyl acetate
	Pyridine
	Pyrrolidine
	Thymol
	Toluene
	Trans-cinnamaldehyde
	Triethylamine

PARTIE 3 : Relations structure-fonction des récepteurs olfactifs

Article 8 - Yu, de March et al. soumis à PNAS (2015)

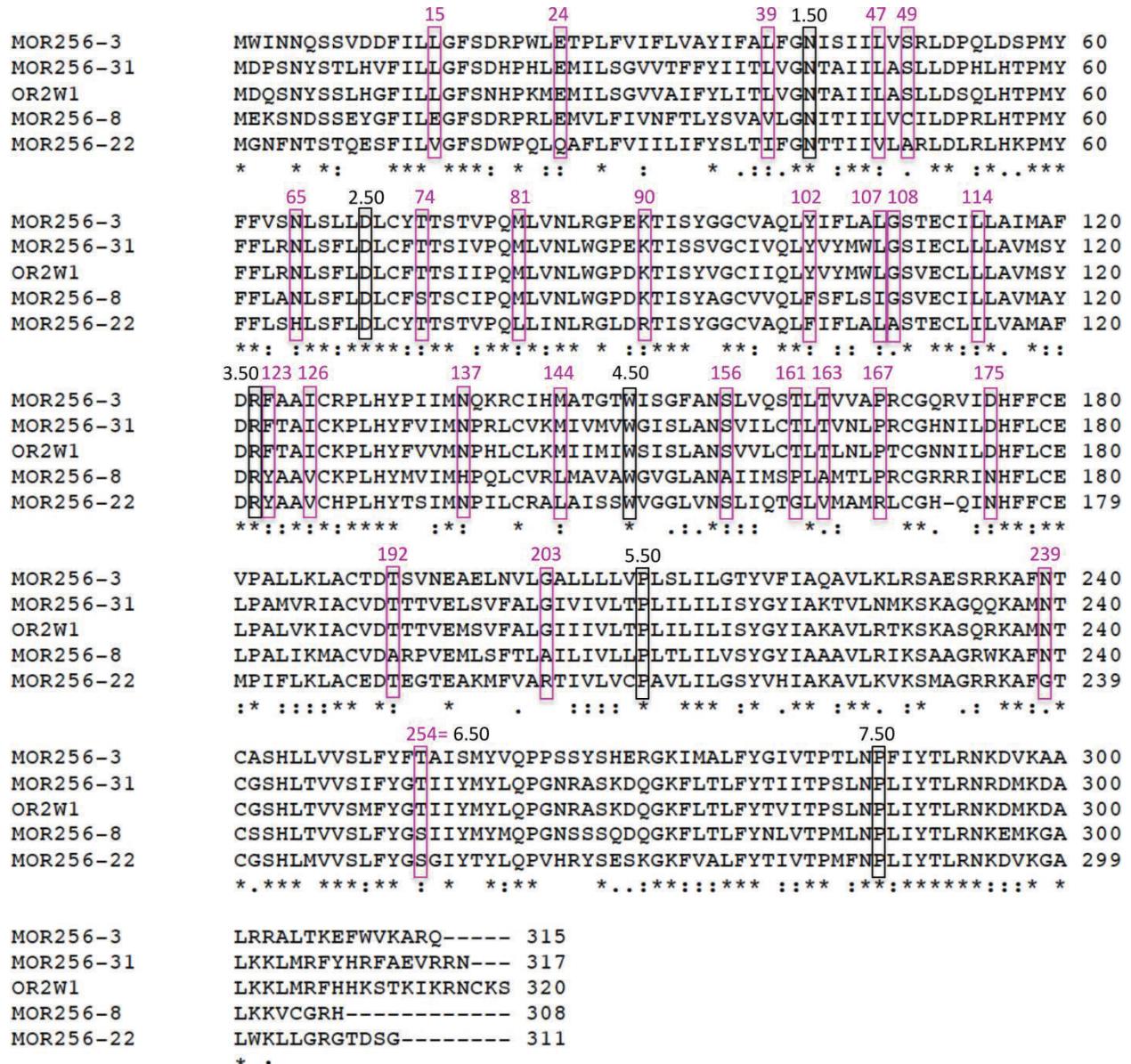


Figure S2. Alignment of the five ORs in the MOR256 subfamily. The Ballesteros-Weinstein notations (1.50 to 7.50) are marked by black rectangles around the corresponding residues. The residues circled in magenta are conserved in broadly responsive MOR256-3, MOR256-31, and hOR2W1, but not in MOR256-8 and/or MOR256-22, which respond weakly to a few odorants.

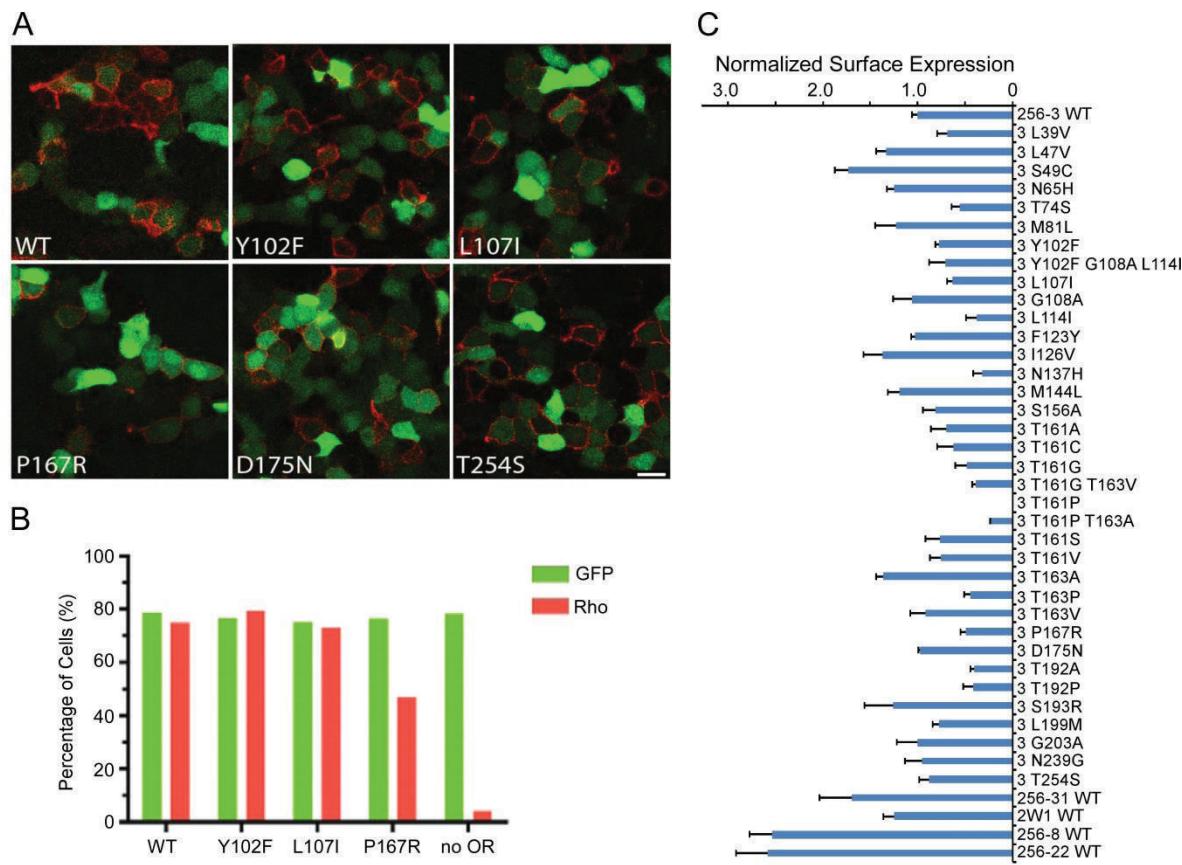


Figure S3. Wild-type ORs and most mutant MOR256-3 ORs are expressed at the cell surface. (A) Hana3A cells were cotransfected with each MOR256-3 OR (wild-type or mutant) and GFP (green) and underwent live cell immunostaining with Rho-antibody (red). (B) For selected ORs, the surface expression was also verified via fluorescence activated cell sorting (FACS). (C) The ratio of red (Rho^+)/green (GFP^+) fluorescence intensity for WT and mutant ORs in live cell immunostaining is used to evaluate the surface expression level. Each data point was averaged from 3-6 plates (mean \pm s.e.m.) and normalized to WT MOR256-3 with the red/green intensity ratio of $53.7\% \pm 3.1\%$.

PARTIE 3 : Relations structure-fonction des récepteurs olfactifs

Article 8 - Yu, de March et al. soumis à PNAS (2015)

	MOR256-3 ORs	Oct	Car	Cou	Benzyl	Allyl	Total
	WT	1.00	0.69	0.54	0.31	0.07	2.60
TM1	L39V	1.50	1.27	1.10	1.02	0.26	5.15
	L47V	0.76	0.50	0.52	0.46	0.12	2.37
	S49C	0.66	0.54	0.60	0.47	0.17	2.43
TM2	N65H	0.83	0.60	0.62	0.46	0.14	2.65
	T74S	1.76	0.99	0.90	0.65	0.11	4.40
	M81L	0.92	0.60	0.56	0.47	0.11	2.66
TM3	Y102F	0.35	0.09	0.05	0.04	0.04	0.58
	Y102F G108A L114I	0.15	0.07	0.06	0.08	0.05	0.41
	L107I	0.42	0.09	0.05	0.06	0.04	0.67
	G108A	0.42	0.23	0.16	0.25	0.09	1.15
	L114I	1.12	0.41	0.29	0.28	0.01	2.10
	F123Y	1.05	0.92	0.81	0.50	0.18	3.46
IC2	I126V	0.73	0.67	0.50	0.31	0.09	2.31
	N137H	2.57	1.46	1.10	0.58	0.14	5.85
TM4	M144L	1.04	0.89	0.78	0.44	0.16	3.32
	S156A	1.04	0.89	0.51	0.21	0.06	2.70
EC2	T161A	0.68	0.30	0.11	0.16	0.01	1.26
	T161C	0.86	0.23	0.21	0.09	0.06	1.44
	T161G	0.49	0.36	0.14	0.30	0.00	1.29
	T161G T163V	0.13	0.05	0.09	0.09	0.11	0.47
	T161P	0.00	0.00	0.04	0.06	0.04	0.15
	T161P T163A	0.33	0.09	0.19	0.23	0.17	1.01
	T161S	1.29	0.49	0.52	0.15	0.06	2.52
	T161V	0.65	0.45	0.16	0.21	0.01	1.48
	T163A	0.59	0.49	0.34	0.33	0.09	1.84
	T163P	0.03	0.04	0.06	0.08	0.07	0.28
	T163V	0.59	0.50	0.14	0.18	0.03	1.43
	P167R	0.82	0.33	0.20	0.12	0.07	1.54
	D175N	1.38	0.94	0.80	0.30	0.20	3.62
TM5	T192A	0.86	0.22	0.12	0.10	0.06	1.36
	T192P	0.19	0.13	0.11	0.11	0.06	0.60
	S193R	0.67	0.32	0.30	0.19	0.03	1.51
	L199M	0.77	0.39	0.62	0.30	0.10	2.18
	G203A	0.97	0.36	0.69	0.45	0.17	2.63
TM6	N239G	1.47	1.15	1.10	0.98	0.33	5.02
	T254S	1.13	0.35	0.46	0.41	0.09	2.44

Table S1. Odorant responses of WT and mutant MOR256-3 ORs. For each OR, the responses to individual odorants (all at 300 µM) were normalized to WT response to 1-octanol and corrected for surface expression. The total response was the sum of the responses to all five odorants (summarized in Fig. 2A). Each mutant OR was tested on the same plate as WT (three repeats for each OR) and the data (with error bars) were also shown in Fig. 4. Oct = 1-octanol, Car = (-) carvone, Cou = coumarin, Benzyl = benzyl acetate, Allyl = allyl phenyl acetate.

Articles 9 et 10 - Activation des récepteurs - vers la déorphanisation computationnelle

Afin de décrypter le code combinatoire lié à la perception olfactive, il est nécessaire de pouvoir identifier l'activation neuronale engendrée par une molécule odorante. Nous avons évoqué précédemment que dans la mesure où un neurone n'exprime qu'un seul type de récepteur, l'activation d'un neurone est assimilable à celle d'un RO. Cette activation est déclenchée par la présence d'une molécule odorante agoniste dans la cavité du récepteur.

Nous avons montré que l'affinité est un paramètre efficace pour discriminer les agonistes des non-agonistes d'un RO. Toutefois, il n'est pas exclu que cette grandeur devienne limitée dans certains cas. En effet, les molécules antagonistes et agonistes inverses se lient à la cavité d'un récepteur mais n'en déclenchent pas l'activation. Dans ces deux cas, le calcul d'une enthalpie libre de liaison est insuffisant pour les discriminer des agonistes car ces ligands montrent une certaine affinité pour le RO étudié. Un mécanisme plus subtil doit être décrit.

Peut-on prédire de façon plus pertinente l'activation d'un neurone ?

Idéalement, la prédiction de l'activation du RO peut être reliée de façon plus directe à l'activation d'un neurone. Au niveau du RO, elle correspond aux changements conformationnels associés au couplage à une protéine G. Des modèles de ce type de mécanisme ont déjà été obtenus sur des RCPG non-olfactifs et ont prouvé leur efficacité.

L'étude en modélisation moléculaire de cette étape nécessite la réalisation de simulations à durée suffisamment longue permettant d'échantillonner le comportement dynamique du récepteur et d'observer le passage d'une conformation dite « inactive » à une autre dite « active » si un agoniste est présent dans sa cavité. Dans une première étude, nous nous focalisons sur l'étude d'un RO considéré comme prototypique d'un RO de mammifère. A travers une approche conjointe entre modélisation moléculaire, analyses bio-informatiques et génie génétique associée à de l'expression *in vitro*, nous identifions les résidus dans les séquences de RO qui gouvernent son activation.

Le modèle du récepteur construit par modélisation est en accord avec les mutations expérimentales. Sa réponse à 5 ligands est à la fois mesurée *in vitro* et calculée. Les résultats théoriques et expérimentaux sont totalement en accord, justifiant le caractère prédictif du modèle.

De façon intrigante, la mutation d'une position de résidu (la position 108) engendre un mutant possédant une haute activité basale (le récepteur est activé sans l'intervention d'un agoniste). Ceci peut être traduit par une plus grande stabilité de l'état activé du récepteur en comparaison à son

Articles 9-10

état inactif. On parle de mutant constitutionnellement actif (CAM). Dans notre approche informatique, le modèle du CAM évolue sans contrainte vers une structure typique de RCPG activé. Cet évènement est une preuve de principe que l'approche computationnelle est suffisamment fiable pour s'apparenter à un microscope computationnel, à travers lequel on peut prédire l'état activé d'un RO.

Cependant, dans le but de s'approcher au maximum de l'obtention d'un outil capable de décrypter le code combinatoire associé aux odeurs, il est nécessaire d'échantillonner des activations de RO qui soient spécifiquement dépendantes d'un agoniste. Dans une deuxième étape, nous réalisons une étude dans laquelle un récepteur est différentiellement activé par un groupe de trois ligands.

*Détermination du caractère agoniste ou non-agoniste par l'observation *in silico* de l'activation d'un RO*

Le large espace chimique des odorants et le nombre important de ROs chez l'humain rendent la combinaison RO/odorant agoniste très vaste. La sélection d'un couple RO/agoniste pertinente est importante pour la suite de l'étude. Idéalement, dans le cadre de la mise au point d'une preuve de principe, le RO doit posséder un agoniste connu expérimentalement. Les couples OR/agoniste hOR7D4/androstenedione et androstanedione sont choisis pour le reste de l'étude sur l'activation des ROs. Un contrôle négatif hOR7D4/non-agoniste ((Z)-2-décenal) est ajouté à l'étude pour s'assurer du caractère sélectif du modèle.

Les simulations de dynamiques moléculaires sur ces systèmes ont permis de reproduire les données expérimentales, c'est-à-dire d'échantillonner des états actifs dans le cas des agonistes et des états inactifs dans le cas du non-agoniste. La prédiction du phénomène d'activation a donc été possible grâce aux techniques de modélisation moléculaire. Sur la base du modèle, des résidus sont suspectés d'être impliqués dans la communication entre la cavité de liaison aux odorants et le site de couplage à la protéine G et donc dans le phénomène d'activation.

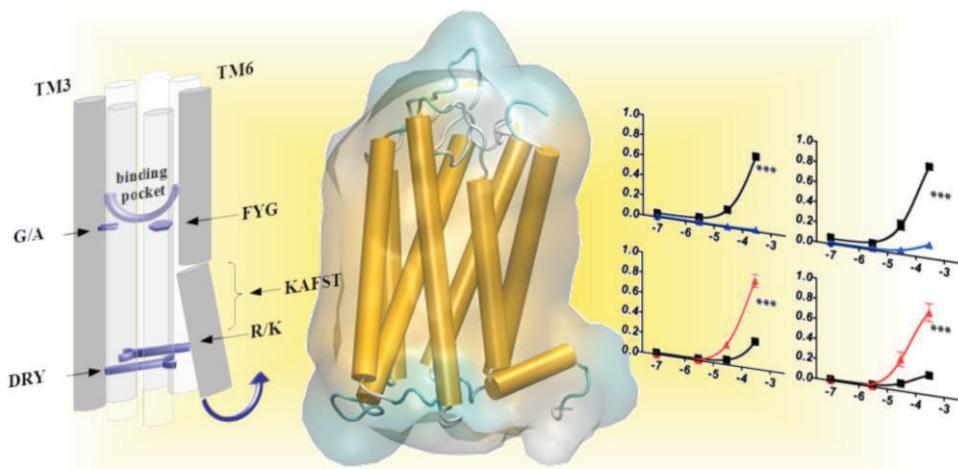
Cette activation échantillonnée par des modèles théoriques est une preuve de principe que la création d'un nez virtuel est maintenant possible.

Dans ces articles, j'ai réalisé la totalité des manipulations de modélisation moléculaire, de bioinformatique structurale ainsi que leurs analyses. Nos collaborateurs des Universités Duke et de UPenn ont fourni les résultats de génie génétique *in vitro*.

Article 9:

Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors

Claire A. de March, Yiqun Yu, Mengjue J. Ni, Kaylin A. Adipietro, Hiroaki Matsunami, Minghong Ma, Jérôme Golebiowski, Journal of the American Chemical Society 137 (2015) 8611-8616



Abstract

Odorant Receptor (OR) genes and proteins represent more than 2% of our genome and 4% of our proteome and constitute the largest sub-group of G Protein-Coupled Receptors (GPCRs). The mechanism underlying OR activation remains poorly understood, as they do not share some of the highly conserved motifs critical for activation of non-olfactory GPCRs. By combining site-directed mutagenesis, heterologous expression, and molecular dynamics simulations that capture the conformational change of constitutively active mutants, we tentatively identified crucial residues for the function of these receptors using the mouse MOR256-3 (Olfr124) as a model. The toggle-switch for sensing agonists involves a highly conserved tyrosine residue in helix VI. The ionic-lock is located between the ‘DRY’ motif in helix III and a positively charged ‘R/K’ residue in helix VI. This study provides an unprecedented model that captures the main mechanisms of odorant receptor activation.

Introduction

The strategy used by mammals to sense odorant molecules is a combinatorial code based on the differential activation of a large family of Odorant Receptors (ORs).[1] One of the major functions of these receptors is to transmit external signals from the environment (odorant molecules) to the nervous system. Furthermore, these proteins are also expressed in non-olfactory tissues, highlighting their role beyond odor detection and potential as drug targets.[2] The OR genes and proteins represent more than 2% of our genome and 4% of our proteome.[3] ORs are seven trans-membrane (TM) helix proteins constituting the largest sub-group of the class-A G protein-coupled receptor (GPCR) family.

GPCRs play critical roles in cellular signal transduction. The initial activation relies on conformational switches between inactive and active states, which depend on both the nature of the receptor and the eventually bound ligand.[4-7] Upon agonist binding, the receptor switches from an inactive to an active form that couples with the intracellular G protein to trigger signal transduction. Inspired from experimental structural data of some GPCRs in active and inactive states, Molecular Dynamics (MD) simulations have been adopted to reveal the atomic-level steps involved in GPCR activation. These models have successfully recapitulated activation of the $\beta 2$ -adrenergic, rhodopsin, muscarinic, and A_{2A} receptors,[6, 8-10] suggesting that this tool is well suited to decipher OR activation. From a mechanistic perspective, GPCR activation is notably associated with the opening of a cleft between the intracellular parts of transmembrane domains 3 and 6 (TM3 and TM6).[11, 12] Several motifs are shown to be important for their activation, e.g., the conserved DRY motif in TM3, the CWxP motif in TM6, and the NPxxY motif in TM7.[7]

ORs have a low sequence identity with other class-A GPCRs. They nonetheless show the same highly conserved motifs within most TM domains, suggesting a conserved general mechanism for their function.[13, 14] ORs also contain some specific motifs, considered as hallmarks for their identification, such as MAYDRYVAICxPLxY in TM3 or KAFSTCxSH in TM6.[15, 16] However, the CWxP motif that plays the role of toggle-switch for GPCR activation is lacking in ORs. Also, although the DRY motif remains highly conserved, the negatively charged residue (E/D in non-olfactory GPCRs) of TM6 facing the DRY motif and involved in the ionic lock between TM3 and TM6 is to date unidentified in ORs. No crystallographic structure of an OR is available and most mechanistic studies rely on the use of molecular modeling.[13]

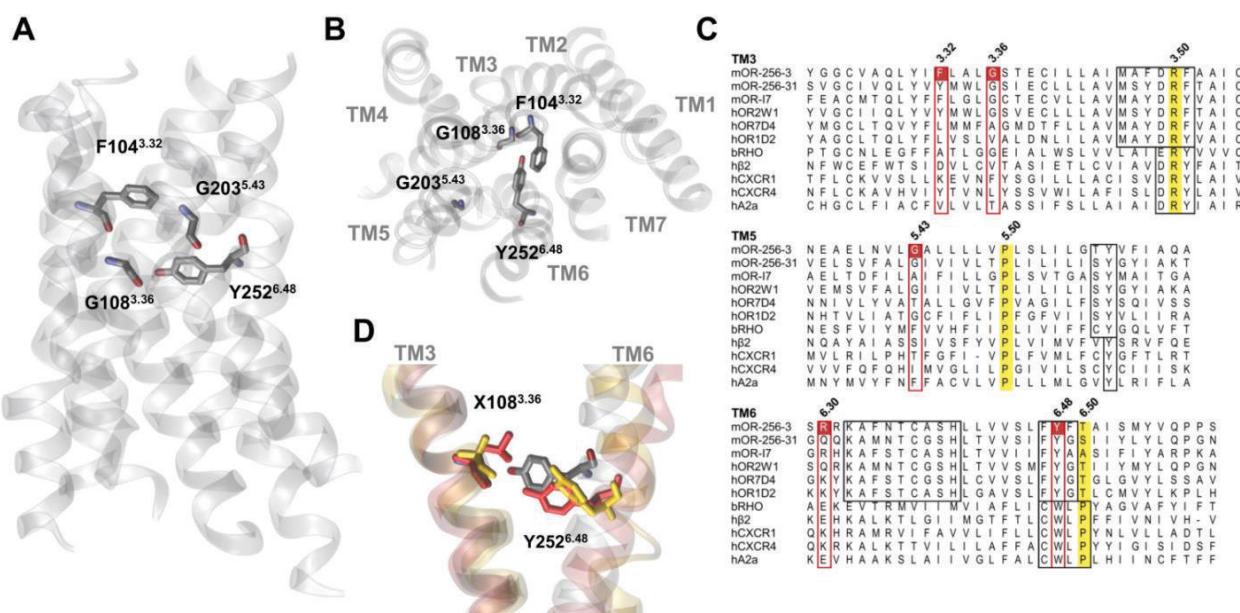


Figure 1. Binding cavity residues that interact with agonists according to the model. (A) & (B) Front view and top view, respectively, of MOR256-3 highlighting selected residues belonging to the binding cavity. (C) Alignments highlighting equivalent roles of certain residues within TM3, TM5 and TM6 in olfactory and non-olfactory GPCRs. The Ballesteros-Weinstein notation is shown for each TM and residues corresponding to the reference position (x.50, with x the TM number) are highlighted in yellow in the alignment. Conserved motifs in ORs and GPCRs are boxed in black. Boxed in red are the residues corresponding to those mutated in this study. (D) The position of Y252^{6.48} shifts as a function of residue X108^{3.36}; wt (*wild type*, G108) is shown in gray, X108= A (G108A) in yellow and X108=L (G108L) in red.

When combined with site-directed mutagenesis data, molecular modeling has led to identification of some specific residues for ligand binding. To date, most studies have focused on the binding cavity, which is consistently made up of residues within TM3, TM5, TM6 and TM7.[17-23] In some ORs, the location of a copper ion as co-factor for detection of sulfur compounds involves residues belonging to the canonical binding site. This supports a conserved activation mechanism, while the metal would

only play a role in ligand affinity,[22, 24] although it does not completely rule out that some ORs function as metalloproteins.[25] The ion-odorant complex might be detected as a single ligand, whose presence may be sensed through a similar mechanism as for all ORs.

Clues about residues potentially involved in the OR activation mechanism were tentatively proposed but they still remain to be assessed by means of *in vitro* experiments and long scale MD simulations.[26-28] Residues involved in the dynamic process that converts an inactive OR structure into an active one are still elusive. This article is a step forward in their identification. The mouse receptor MOR256-3 (also named Olfr124 or SR1), a broadly-tuned receptor,[29] is the focus of the current study by a joint approach that combines molecular modeling, site-directed mutagenesis and heterologous expression. The MOR256-3 sequence contains the hallmarks of mammalian ORs, with typical highly conserved motifs in all TMs (Figure S1 and Table S1). Their conservations were assessed by a thorough sequence analysis on 396 human and 1111 mouse ORs. We provide a body of evidence for the functional role of several of these motifs within OR sequence. Based on an experimental observation of mutant ORs with either increased basal activity (ligand-independent receptor activation) or locked into a constitutively active state, we have built a structural model that captures this active state while the wild-type (*wt*) OR stays in an inactive form. This offers the opportunity to decipher the strategy used by ORs to detect agonists without being subjected to the difficult task of finding the accurate position of the ligand within the binding site. We show that in ORs, the highly conserved Y residue of the FYG motif in TM6 acts as the toggle-switch. Also, the ionic-lock involves the D and R residues of the DRY motif in TM3 and a conserved positive residue in TM6. This is the first report of a homology OR model that evolves between active and inactive states.

Results

Residues within the binding cavity control ligand specificity and basal activity

To gain insights into the OR activation mechanism, a 3D atomic computational model of MOR256-3 was built by homology modeling using an alignment and a multi-template approach shown to be consistent with experimental affinity data on OR-ligand pairs.[13, 30] Most amino acid residues identified as belonging to the binding site (made up of residues from the upper parts towards the extracellular side) of the helices of TM3, TM5, TM6 and TM7) are consistent with previous studies (ref. [31] and references therein). Residues $104^{3.32}$, $108^{3.36}$, $203^{5.43}$, and $252^{6.48}$ are notably pointing into the binding site, as shown in Figure 1, where the superscript numbers are the Ballesteros-Weinstein notation in the alignment (Figures S1 and S2). These residues have been identified in ligand recognition in many studied ORs so far, including copper-mediated ORs.[19-22]

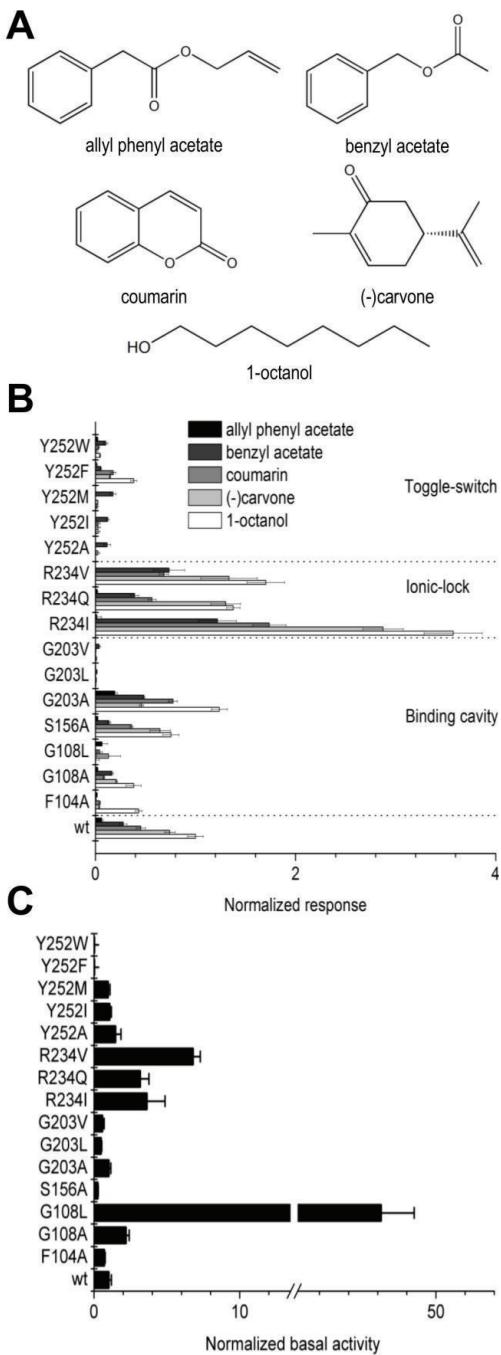


Figure 2. Odorant-evoked responses and basal activities of MOR256-3 receptors. (A) Structures of odorants tested in this study. (B) MOR256-3 wild-type (*wt*) and mutant responses to odorants. All data are normalized to the *wt* response to 1-octanol. (C) Basal activities of the same receptors, normalized to that of the *wt* OR. The five odorants are ranked based on the responses of the *wt* receptor from largest to smallest. Mutations include residues presumably involved in the binding cavity, in the ionic-lock, and in the toggle-switch. Data for each OR are averaged from three repeats on the same 96-well plate except for WT, averaged from 12 plates with three repeats on each (mean \pm s.e.m.).

Notice that residues 104^{3.32}, 108^{3.36}, and 252^{6.48} make consensus contacts with ligands across class-A GPCRs.[14] To assess the functional role of residues of interest in MOR256-3 and its mutants, both basal activity and responses elicited by a set of chemically diverse odorants were measured (Figure 2; for dose-response curves, see Figure S3).

Five odorants ranging from strong to weak agonists were selected to cover agonists with a range of potency (Figure 2A). As predicted by the model, the F104A^{3.32} mutant indeed displays altered agonist recognition by modifying the selectivity of the receptor *in vitro*. In this mutant OR, the response to all agonists tested is generally decreased and the receptor only responds moderately to octanol, which is the only odorant that lacks a π-cloud.

Response of this mutant to odorants suggests that F104^{3.32} contributes to stabilizing bound ligands through an interaction between its aromatic cycle and double bonds present in odorants. The G203A^{5.43} mutant presents a short hydrophobic side-chain that does not dramatically modify accessibility to the cavity but is likely to contribute to *van der Waals* contacts that slightly increase the response to odorants. Upon increase of the side-chain size (G203V or G203L), the mutants do not respond to odorants anymore highlighting that this position is also within the binding cavity. In accordance with the model, the S156A^{4.57} mutation has no influence on odorant recognition, as its side-chain is located outside the binding cavity. This model also identifies other residues contributing to receptor selectivity, as reported by Yu *et al.*[32] Remarkably, without ligand stimulation, G108A and G108L display unique behaviors. G108A shows a basal activity that is twice higher as the *wt*, while G108L has a basal activity ~45 times higher (Figure 2C). Interestingly, similar modulation upon mutating G108 has been observed in MOR256-31 (Figure S4) confirming that this effect is not specific to our model MOR256-3.

When stimulated with each odorant, G108A is still able to discriminate between weak and strong agonists but with much weaker responses than the *wt*. G108L, however, is virtually unresponsive to agonists (Figure 2B). These data suggest that the MOR256-3 G108A mutant favors the active state while the G108L mutant is locked into a constitutively active state.

The FYG motif in TM6 is associated with the toggle-switch for sensing agonists

Position 108^{3.36} is not highly conserved among ORs but it is represented in more than 85% ORs by a small residue (see Table S1), suggesting that this part of the binding cavity must be accessible to odorants. In the G108A or G108L mutants, the side chain is pointing towards Y252^{6.48} where a Y/F residue is conserved in more than 92% in human and mouse ORs. These two residues are reported to form a ligand-binding cradle across class-A GPCRs.[14] Interestingly, Y/F^{6.48} is aligned with the tryptophan residue W^{6.48} of the highly conserved CWxP motif in non-olfactory class-A GPCRs (Figure 1C), reported as a toggle-switch for receptor activation.[33] Here, the side chain of residue 108 in the

mutants is likely to play the role of an artificial agonist which interacts with the side-chain of Y/F252^{6.48} (Figure 1D). Accordingly, we tested substitutions at position 252 with several different amino-acids. MOR256-3 Y252A, Y252I and Y252M mutants, although expressed at the cell surface (Figure S5), do not exhibit any statistically significant *in vitro* response upon odorant stimulation (Figure 2B and Figure S3; note that all odorant responses are corrected for the expression efficiency). These data are consistent with the role of Y/F252^{6.48} as a toggle-switch that triggers activation of the receptor upon agonist binding.[20, 34] The difference between tyrosine and phenylalanine was also investigated. Consistent with an F conserved in ~25% mammalian ORs at position 252, the Y252F mutant shows responses to some odorants. Its responsiveness is however decreased by 70% compared with the *wt*. Contrasting with non-olfactory GPCRs, when position 6.48 is a tryptophan residue (not found in native ORs), the OR becomes almost non-responsive (~5% of the *wt* response) (Figure 2A).

Molecular dynamics simulations model active and inactive states and identify the ionic-lock residues

GPCR activation is associated with a conformational change involving the ionic-lock between TM3 and TM6.[12] MD simulations performed on models of the *wt*, G108A and G108L free of agonists reveal interesting structural features. In the *wt*, one systematically observes a structure showing the hallmarks of an inactive state, with the bottom (intracellular side) of the helix of TM6 close to that of TM3.

In the G108A and G108L models, the simulations converge towards a structure where TM6 has shifted from its initial position and moved ~8Å outward (Figure 3), which is the signature of GPCR activation.[12] For these two mutants, all four independent simulations converge to an alternative model of the MOR256-3 receptor that closely resembles crystallographic structures of GPCRs in an active state (Figure 3A, B).

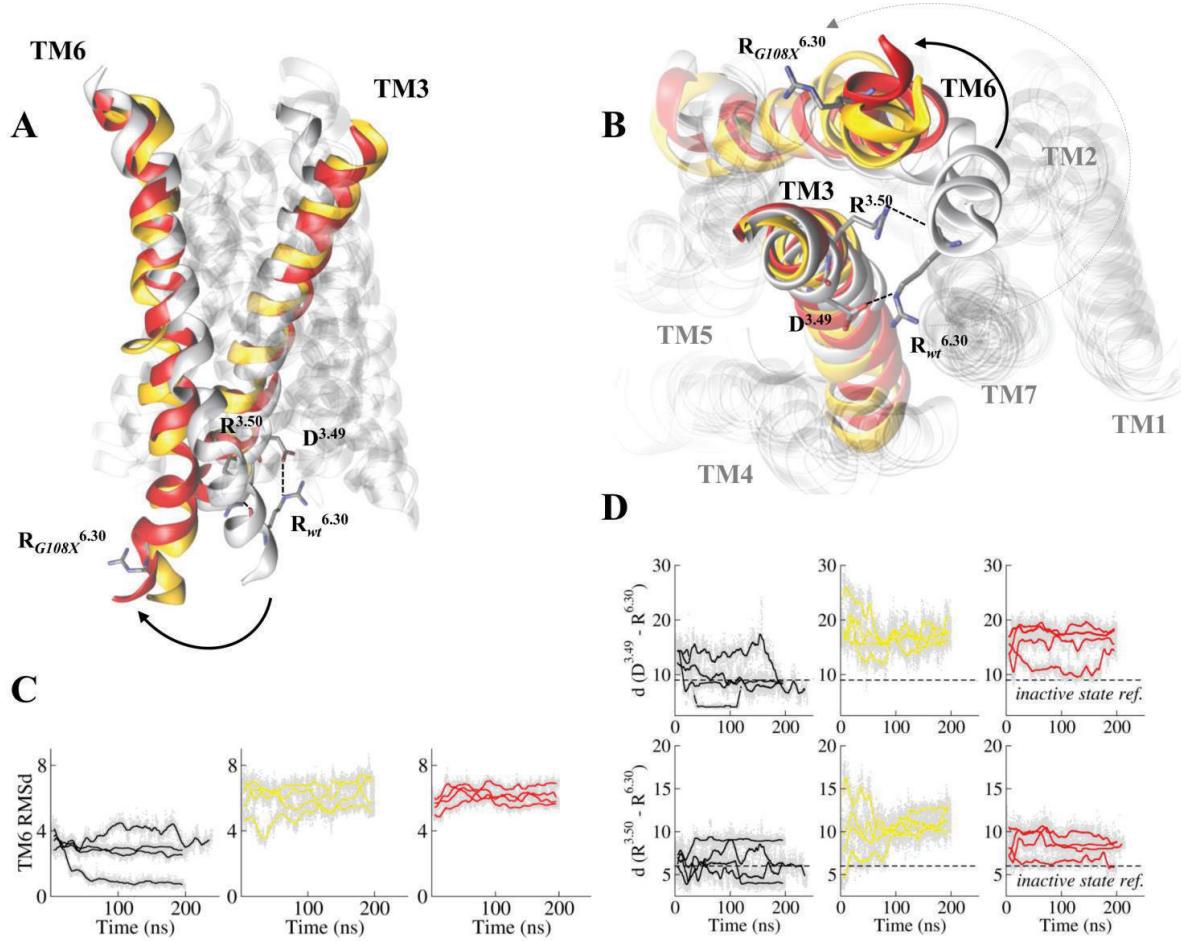


Figure 3. MOR256-3 *wt* systematically reports an inactive state, while mutations at position 108 evolve toward active states. (A) Comparison between typical structures of MOR256-3 *wt* (white), G108A (yellow) and G108L (red) mutants. (B) Structures of the mutants have the intracellular part of TM6 shifting outward while that of the *wt* is close to other TMs. (C) The Root Mean Square deviation (RMSd) of TM6 (residues 234 to 253) with respect to its reference position in the *wt* structure reveals a systematic shift in the mutants (RMSd \approx 6 Å) but not in the *wt* (RMSd \approx 3 Å). (D) The ionic-lock between R^{6.30} and D^{3.49} as well as between R^{6.30} and R^{3.50} is closed in the *wt* ($d(O_D^{3.49}...N_R^{6.30}) \sim 9$ Å and $d(N_R^{3.50}...O_R^{6.30}) \sim 6$ Å) while it is open in the mutants (the distance becomes ~ 17 Å and 9 Å, for $d(O_D^{3.49}...N_R^{6.30})$ and $d(N_R^{3.50}...O_R^{6.30})$, respectively).

The *wt* model systematically presents a double interaction between the D^{3.49} and R^{3.50} (conserved at 98% and 88%, respectively) of the DRY motif in TM3 on one part, and the R^{6.30} side-chain and backbone in TM6 on the other part. A positive residue (R/K) at this position in TM6 is highly conserved in ORs (more than 75%, see Table S1) and aligned with the residue involved in the ionic-lock in non-olfactory GPCRs (Figure 1C). The interaction between TM6 and TM3 at the ionic-lock involves the side-chain and the backbone of R^{6.30} and the side chains of D^{3.49} and R^{3.50}, respectively. These interactions are observed during three out of the four simulations of the *wt* system, as shown in Figure 3C, D. The four independent simulations performed for each G108X mutant systematically

report a typical structure where the interactions between TM3 and TM6 are broken. Very early in the equilibration phases of the mutant receptors, TM6 shifts outward relative to TM3 (see arrows in Figure 2A, B) while it stays in its initial position in the *wt*. An analysis of the root mean square deviation of TM6 heavy atoms with respect to their average position in the *wt* reveals a large structural drift of TM6 while the rest of the edifice remains similar to the starting structure (Figure S6).

This conformational switch is in all cases associated with a break of the hydrogen bond between D^{3.49} and R^{6.30}. The distance between the closest H-bond donor and acceptor atoms within these residues is ~17 Å in the mutants while it is ~9 Å in the *wt*. Such a distance evolution (an increase of 8 Å) is consistent with those measured in experimental structures.[7]

The second interaction between the R^{3.50} side-chain and the R^{6.30} backbone oxygen atom is also largely weakened, with a distance 3 Å larger in the mutants compared with the *wt*, in line with experimental data on rhodopsin.[12] In the G108X mutants, the R^{6.30} side-chain has shifted towards the intracellular part of the receptor and is solvated by bulk water. D^{3.49} forms a hydrogen bond with either Y132^{IL3} or/and R^{3.50} which has also broken its interaction with TM6, consistent with MD simulations performed on X-ray structures of GPCRs.[8, 35] Experimentally, when position 6.40 is modified to a non-polar residue that prevents any interaction with TM3 (R234I or R234V mutants), one observes an increase of the response to odorants together with high basal activity, consistent with a shift of the conformational equilibrium of the receptor towards the active state. In the R234Q mutant, the charged D^{TM3} – R^{TM6} ionic-lock is altered from an ionic interaction into a hydrogen bond and the receptor now exhibits similar, although weaker, increase of basal activity and responsiveness (Figure 2B, C). In all cases, the ranking of the odorants remains mostly unchanged, confirming that this modification is taking place at residues involved in OR activation rather than recognition of odorants.

Discussion

A model for OR activation

The mammalian olfactory system uses a combinatorial strategy based on a large family of ORs to sense odorant compounds. As for all GPCRs, when a receptor is activated by an agonist, the coupling to the G protein occurs, while in its inactive state, a receptor does not trigger the biochemical cascade leading to neuron membrane depolarization. At the atomic-level, although the mechanism is becoming more and more precisely understood for non-olfactory GPCRs, that of ORs remains elusive. In this study, using *in vitro* observations of mouse OR mutants in a constitutively active state, we provide insights into the way specific amino acid residues lead to the activation of an OR.

Several motifs are highly conserved in class-A GPCRs. The three-residue E/DRY motif in TM3 is involved in a so-called ionic-lock with a residue in TM6. In ORs, this DRY motif is also highly conserved although the D residue is predominant with respect to the R (Table S1), contrary to what is seen in other non-olfactory GPCRs. Indeed, in non-olfactory GPCRs, TM3 interacts with TM6 through the positively charged R^{3.50} of the DRY motif and a negatively charged E^{6.30} (as observed in rhodopsin,[36] β1[37] and β2-adrenergic,[38] human D3-dopamine,[39] human H1-histamine,[40] human M2-muscarinic,[41] and A2A adenosine receptors).[42] In the inactive state structure, this ionic-lock is closed while in the active state it is open. Chemokine receptors CXCR4 and CXCR1 stand as an exception since they exhibit a positively charged residue at position 6.30.[43, 44] In the crystal structure of these molecules, no ionic lock is observed but instead a charge dipole interaction exists involving the R^{3.50} side chain and the backbone of TM6 at residue in position 6.30. The case of ORs appears related to those of CXCRs. Our model is indeed in line with this structural feature with a strong hydrogen bond between R^{3.50} and the backbone of R^{6.30}. In addition, we show that the D^{3.49} residue is engaged in a hydrogen bond with an arginine (R) residue in TM6 to stabilize the TM3-TM6 interaction, justifying its high degree of conservation in ORs. The R residue of the DRY motif is also engaged in a strong interaction with the backbone of residue 6.30. Considering MOR256-3 as a prototype of mammalian ORs (because of its conserved residues with all other mammalian ORs in this motif), a double ionic-lock seems prevalent between TM3 and TM6 in this family.

It is very likely that the active state of ORs is highly similar to that of class-A GPCRs as revealed by X-Ray crystallography. Our molecular models are in full accordance with a conserved mechanism of activation. In the multiple MD simulations we recover—without any constraint—a cleft between TM3 and TM6 in OR mutants with a higher basal activity or in a constitutively active state. Although MD simulations have already reported active and inactive GPCR structures, they were all based on experimental data.[4, 6, 8-10]

The model of an active OR was made possible by the unique behavior associated to mutations at position 108^{3.36}. Notice that a small perturbation at position 107^{3.35} in MOR256-8 strongly increases the basal activity and affect responsiveness of the mutant OR, confirming that this part of the receptor is crucial for activation.[32] Position 108^{3.36} in TM3 is associated with a small residue that may allow space for agonist binding deep into the pocket. The nature of this residue affects responsiveness of the receptor, as previously reported for a hOR1A2 A108G mutant.[18] Position 108 faces Y252^{6.48}, which is highly conserved in ORs and interestingly aligned with the toggle switch residue (W^{6.48} of the ‘CWxP’ motif) found in non-olfactory GPCRs. Consequently, Y252^{6.48} can be considered to be the toggle-switch in ORs. This residue has been speculated to be involved in OR activation by agonists but has not been clearly assessed.[19, 20] Our current data confirm this

hypothesis, similar to what has been shown in the A3 adenosine receptor.[34] Site-directed mutations suggest that the toggle-switch should share physico-chemical properties with the associated OR ligands. Airborne odorants are more hydrophobic than non-olfactory GPCR ligands. Accordingly, based on sequence analysis, the transmission switch for agonists within the cavity of an OR is an aromatic residue (Y/F) at position 6.48. The tryptophan cycle cannot play the role of toggle-switch in ORs, because of either too-large hydrophilicity or a too-bulky character. In the first case, the interaction with agonists wouldn't be favored. Note that the contribution of residue 6.48 to the free energy of binding is computed to be important when agonists are bound to hOR1G1.[30, 45]

Once the agonist is bound within the OR cavity, the activation process propagates by creating a drift of TM6 which will, as for other GPCRs, open a cleft at the intra-cellular part of the bundle to favor G protein coupling. Interestingly, the part of TM6 that moves outward involves the highly conserved KAFSTCxSH motif, consistent with both the role of the serine residue (S) in the change from the active to the inactive conformation and more generally the contribution of this motif to the receptor conformation.[46]

The multidisciplinary approach used here is promising in elucidating the activation process of receptors with unknown experimental structures. In this article, we focused on residues belonging to TM3, TM5 and TM6. These residues studied in MOR256-3 are highly conserved in human and mouse ORs. Since MOR256-3 is broadly tuned, there is a possibility that the proposed mechanism only applies to broadly-tuned receptors. However, the identified residues are conserved in both broadly-tuned and narrowly-tuned ORs suggesting that the mechanism may apply to all ORs independent of their tuning properties. Further investigations with narrowly-tuned receptors are required to distinguish these two possibilities

Other parts of the OR are also surely important for OR activation by fulfilling the network of amino acids involved in the process from ligand binding to G protein coupling. This is for example the case of the conserved NPxxY motif within TM7.[14, 47, 48] This approach nevertheless provides a fruitful working model for OR activation based on site-directed mutagenesis and molecular dynamics simulations, which is of high importance for predicting olfactory sensory neuron responses upon ligand stimulation.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] C. Flegel, S. Manteniotis, S. Ostholt, H. Hatt, G. Gisselmann, Expression Profile of Ectopic Olfactory Receptors Determined by Deep Sequencing, *PLoS ONE* 8 (2013) e55368.
- [3] Y. Niimura, Olfactory Receptor Multigene Family in Vertebrates: From the Viewpoint of Evolutionary Genomics, *Curr Genomics* 13 (2012) 103-114.
- [4] R.O. Dror, D.H. Arlow, P. Maragakis, T.J. Mildorf, A.C. Pan, H. Xu, D.W. Borhani, D.E. Shaw, Activation mechanism of the β 2-adrenergic receptor, *Proc Natl Acad Sci U S A* 108 (2011) 18684-18689.
- [5] A.C. Kruse, A.M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hubner, E. Pardon, C. Valant, P.M. Sexton *et al*, Activation and allosteric modulation of a muscarinic acetylcholine receptor, *Nature* 504 (2013) 101-106.
- [6] R. Nygaard, Y. Zou, Ron O. Dror, Thomas J. Mildorf, Daniel H. Arlow, A. Manglik, Albert C. Pan, Corey W. Liu, Juan J. Fung, Michael P. Bokoch *et al*, The Dynamic Process of β 2-Adrenergic Receptor Activation, *Cell* 152 (2013) 532-542.
- [7] B. Trzaskowski, D. Latek, S. Yuan, U. Ghoshdastider, A. Debinski, S. Filipek, Action of molecular switches in GPCRs--theoretical and experimental studies, *Curr Med Chem* 19 (2012) 1090-1109.
- [8] E.N. Laricheva, K. Arora, J.L. Knight, C.L. Brooks, Deconstructing Activation Events in Rhodopsin, *J Am Chem Soc* 135 (2013) 10906-10909.
- [9] Y. Miao, S.E. Nichols, P.M. Gasper, V.T. Metzger, J.A. McCammon, Activation and dynamic network of the M2 muscarinic receptor, *Proc Natl Acad Sci U S A* 110 (2013) 10982-10987.
- [10] J. Li, A.L. Jonsson, T. Beuming, J.C. Shelley, G.A. Voth, Ligand-Dependent Activation and Deactivation of the Human Adenosine A2A Receptor, *J Am Chem Soc* 135 (2013) 8749-8759.
- [11] U. Gether, B.K. Kobilka, G Protein-coupled Receptors: II. MECHANISM OF AGONIST ACTIVATION, *J Biol Chem* 273 (1998) 17979-17982.
- [12] C. Altenbach, A.K. Kusnetzow, O.P. Ernst, K.P. Hofmann, W.L. Hubbell, High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation, *Proc Natl Acad Sci U S A* 105 (2008) 7439-7444.
- [13] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, in: C.J. Crasto (Eds.), Olfactory Receptors, Molecular Modelling of Odorant/Olfactory Receptor Complexes, New York, 2013, pp. 53-65.
- [14] A.J. Venkatakrishnan, X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, M.M. Babu, Molecular signatures of G-protein-coupled receptors, *Nature* 494 (2013) 185-194.
- [15] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, *Genome Biol* 2 (2001) 1-12.
- [16] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human–mouse comparisons, *Protein Sci* 13 (2004) 240-254.
- [17] T. Abaffy, A. Malhotra, C.W. Luetje, The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor: A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES, *J Biol Chem* 282 (2007) 1216-1224.
- [18] K. Schmiedeberg, E. Shirokova, H.-P. Weber, B. Schilling, W. Meyerhof, D. Krautwurst, Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2, *J Struct Biol* 159 (2007) 400-412.

- [19] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, H. Pick, The Mouse Eugenol Odorant Receptor: Structural and Functional Plasticity of a Broadly Tuned Odorant Binding Pocket, *Biochemistry* 50 (2011) 843-853.
- [20] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural Basis for a Broad But Selective Ligand Spectrum of a Mouse Olfactory Receptor: Mapping the Odorant-Binding Site, *J Neurosci* 25 (2005) 1806-1815.
- [21] L. Gelis, S. Wolf, H. Hatt, E. Neuhaus, K. Gerwert, Prediction of a Ligand-Binding Niche within a Human Olfactory Receptor by Combining Site-Directed Mutagenesis with Dynamic Homology Modeling, *Angew. Chem. Int. Ed. Engl.* 51 (2012) 1274-1278.
- [22] S. Sekharan, Mehmed Z. Ertem, H. Zhuang, E. Block, H. Matsunami, R. Zhang, Jennifer N. Wei, Y. Pan, Victor S. Batista, QM/MM Model of the Mouse Olfactory Receptor MOR244-3 Validated by Site-Directed Mutagenesis Experiments, *Biophys J* 107 (2014) L5-L8.
- [23] C.A. de March, S.-K. Kim, S. Antonczak, W.A. Goddard, J. Golebiowski, G Protein-coupled odorant receptors: From sequence to structure, *Protein Sci* (2015) n/a-n/a.
- [24] X. Duan, E. Block, Z. Li, T. Connelly, J. Zhang, Z. Huang, X. Su, Y. Pan, L. Wu, Q. Chi, Crucial role of copper in detection of metal-coordinating odorants, *Proc Natl Acad Sci U S A* 109 (2012) 3492-3497.
- [25] J. Wang, Z.A. Luthey-Schulten, K.S. Suslick, Is the olfactory receptor a metalloprotein?, *Proc Natl Acad Sci U S A* 100 (2003) 3035-3039.
- [26] S.-K. Kim, W. Goddard, III, Predicted 3D structures of olfactory receptors with details of odorant binding to OR1G1, *J Comput Aided Mol Des* 28 (2014) 1-16.
- [27] P.C. Lai, M.S. Singer, C.J. Crasto, Structural Activation Pathways from Dynamic Olfactory Receptor–Odorant Interactions, *Chem Senses* 30 (2005) 781-792.
- [28] P.C. Lai, B. Guida, J. Shi, C.J. Crasto, Preferential Binding of an Odor Within Olfactory Receptors: A Precursor to Receptor Activation, *Chem Senses* 39 (2014) 107-123.
- [29] X. Grosmaire, S.H. Fuss, A.C. Lee, K.A. Adipietro, H. Matsunami, P. Mombaerts, M. Ma, SR1, a Mouse Odorant Receptor with an Unusually Broad Response Profile, *J Neurosci* 29 (2009) 14545-14552.
- [30] J. Topin, C.A. de March, L. Charlier, C. Ronin, S. Antonczak, J. Golebiowski, Discrimination between Olfactory Receptor Agonists and Non-agonists, *Chem Eur J* 20 (2014) 10227-10230.
- [31] C.A. de March, J. Golebiowski, A computational microscope focused on the sense of smell, *Biochimie* 107 (2014) 3-10.
- [32] Y. Yu, C.A. de March, M.J. Ni, K.A. Adipietro, J. Golebiowski, H. Matsunami, M. Ma, Low Activation Threshold Combined with Permissive Binding Pocket Underlies Broad Tuning of Mammalian Odorant Receptors, (submitted)
- [33] X. Deupi, J. Standfuss, Structural insights into agonist-induced activation of G-protein-coupled receptors, *Curr Opin Chem Biol* 21 (2011) 541-551.
- [34] Z.-G. Gao, A. Chen, D. Barak, S.-K. Kim, C.E. Müller, K.A. Jacobson, Identification by Site-directed Mutagenesis of Residues Involved in Ligand Recognition and Activation of the Human A3 Adenosine Receptor, *J Biol Chem* 277 (2002) 19056-19063.
- [35] R.O. Dror, D.H. Arlow, D.W. Borhani, M.Ø. Jensen, S. Piana, D.E. Shaw, Identification of two distinct inactive conformations of the β 2-adrenergic receptor reconciles structural and biochemical observations, *Proc Natl Acad Sci U S A* 106 (2009) 4689-4694.
- [36] T. Okada, M. Sugihara, A.N. Bondar, M. Elstner, P. Entel, V. Buss, The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure, *J Mol Biol* 342 (2004) 571-583.

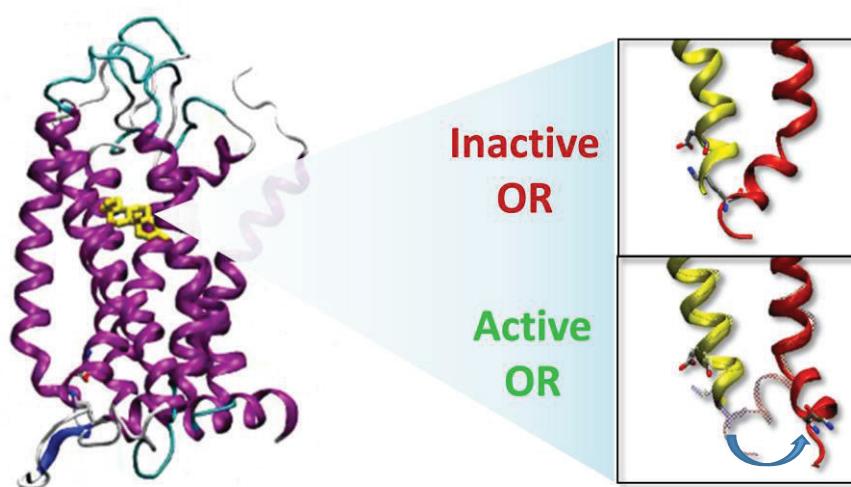
Article 9 – de March et al. JACS 137 (2015) 8611-8616

- [37] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, G.F. Schertler, Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature* 454 (2008) 486-491.
- [38] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka *et al*, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science* 318 (2007) 1258-1265.
- [39] E.Y. Chien, W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov *et al*, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, *Science* 330 (2010) 1091-1095.
- [40] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han *et al*, Structure of the human histamine H1 receptor complex with doxepin, *Nature* 475 (2011) 65-70.
- [41] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga *et al*, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, *Nature* 482 (2012) 547-551.
- [42] G. Lebon, T. Warne, P.C. Edwards, K. Bennett, C.J. Langmead, A.G. Leslie, C.G. Tate, Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation, *Nature* 474 (2011) 521-525.
- [43] S.H. Park, B.B. Das, F. Casagrande, Y. Tian, H.J. Nothnagel, M. Chu, H. Kiefer, K. Maier, A.A. De Angelis, F.M. Marassi *et al*, Structure of the chemokine receptor CXCR1 in phospholipid bilayers, *Nature* 491 (2012) 779-783.
- [44] B. Wu, E.Y.T. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi *et al*, Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, *Science* 330 (2010) 1066-1071.
- [45] L. Charlier, J. Topin, C. Ronin, S.-K. Kim, W. Goddard, R. Efremov, J. Golebiowski, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case, *Cell Mol Life Sci* (2012) 1-9.
- [46] A. Kato, S. Katada, K. Touhara, Amino acids involved in conformational dynamics and G protein coupling of an odorant receptor: targeting gain-of-function mutation, *J Neurochem* 107 (2008) 1261-1270.
- [47] K.J. Kohlhoff, D. Shukla, M. Lawrenz, G.R. Bowman, D.E. Konerding, D. Belov, R.B. Altman, V.S. Pande, Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways, *Nat Chem* 6 (2014) 15-21.
- [48] M. Audet, M. Bouvier, Restructuring G-Protein- Coupled Receptor Activation, *Cell* 151 (2012) 14-23.

Article 10:

Molecular dynamics predicts activation of a human Odorant Receptor upon agonist stimulation

Claire A. de March, Gleb Novikov, Elise Brugera, Hiroaki Matsunami, Jérôme Golebiowski, en préparation



Introduction

Our sense of smell is triggered by the activation of odorant receptors expressed by our olfactory sensory neurons. The consensus is that the coding of an odor by our brain relies on a so-called combinatorial code of OR activation.[1-3] ORs belong to the class A G protein-coupled receptors, a family of seven transmembrane (7TM) proteins responsible for transmitting signals between cells. As all other class A GPCRs, ORs can be activated by agonists, blocked by inverse agonists, or eventually be unaffected in front of neutral antagonists.[4, 5] Upon agonist stimulation, the receptor switches from a predominantly inactive to a predominantly active form. At the atomic level, these states mostly differ at the intracellular part of the transmembrane (TM) domain, which will be coupled to the G protein.[6, 7]

In ORs, since no experimental structure has been elucidated, the mechanism of activation is mostly investigated by means of joint approaches combining molecular modeling inspired from existing class A GPCR structures and *in vitro* site-directed mutagenesis data, see [4, 8] for reviews. Although highly variable residues within the binding site define the binding properties, some conserved residues are involved in receptor activation. At the crossing between TM3 and TM6, some residues form a cradle for the ligand and participate to the cross-talking between the binding site and the intracellular part of the receptor. The function of GPCRs is maintained by highly conserved motifs within their sequences. Although in non-olfactory class A GPCR a CWxP motif is prevalent in TM6, OR exhibit a specific FYG motif in which the second residue is mandatory for activation.[9] Modifications within residues in TM3 facing this motif strongly alter the OR function, with modification of either the basal activity,[9] or the recognition spectrum.[3] Of course other residues surely contribute to defining the network involved in OR activation. The question of the ability of molecular modeling to predict OR activation upon agonist stimulation remains open.

Although many ORs have a relatively large recognition spectrum, OR7D4 stands as an exception. This receptor was shown to be mostly responsible for our detection of androstenone and androstadienone, while many other odorants are unable to trigger activation.[10] OR7D4, indeed, acts as a lock-and-key GPCR that shows a narrow spectrum of odorant recognition. Its sequence shares all the specific motifs with other ORs and is used in this study to gain information on the role of some residues in the spreading of the activation signal from the binding cavity to the G protein coupling site. Using a combination of molecular dynamics based on a homology model and *in vitro* site-directed mutagenesis, we assess the quality of the model, able to recover active / inactive forms when bound to an agonist / neutral antagonist. In the FYG motif of TM6, we show that although the two first residues (FY) must be aromatic (F or Y) to trigger activation by an agonist, the third one (G) can be more variable but should remain tiny.

Results

Binding site residues influence receptor response

The structure of the receptor exhibits the hallmarks of class A GPCRs. Two agonists (androstadienone and androstenone) and a neutral antagonist ((Z)-2-decenal) are docked into the apo receptor prior to microsecond timescale molecular dynamics simulations (*vide infra*). The binding cavity is made up of residues belonging to TM3, TM5, TM6 and TM7. The model predicts a larger affinity of the receptor for the two agonists with respect to the inverse antagonist (Fig. 1), suggesting that agonists must reach a certain affinity to trigger activation.

From a structural point of view, the model predicts several contacts between androstadienone/androstenone and residues of the TM cited above. The list of residues in contact is in line with those identified in other systems.[11] The small side-chain of residue A202 is notably allowing space for the binding of androstadienone/androstenone, consistent with *in vitro* data showing that A202F mutation actually kills the response of the receptor upon ligand stimulation (Fig. 1). Notice that in no case the WT or the mutant ORs showed response upon decenal stimulation although they have good expression levels, (data not shown).

F251, Y252 and G253 are involved in the highly conserved FYG motif of TM6, shown to act as the transmission switch in ORs, equivalent to CWxP in non-olfactory GPCRs.[9] Accordingly, as shown in a previous study, the Y252A is not responsive to agonists *in vitro* anymore (Fig. 1).[9] Here we investigate the effect of mutating the F251 and G253 residues. Similarly to position 252, the F251Y mutant shows unchanged response upon agonist stimulation while the F251A mutant becomes totally unresponsive. The G253A shows a weaker response than the WT while the G253L mutant exhibits a very weak response. This suggests that, within the FYG motif, the two first positions must be aromatic residues, while the third one must be conserved as a small residue, consistent with a residue conservation analysis on mouse and human ORs. In these ORs, positions 251 and 252 are conserved as aromatic residues at more than 79% and position 253 is conserved as G, V, and S with percentages of 65, 11, and 5%, respectively.

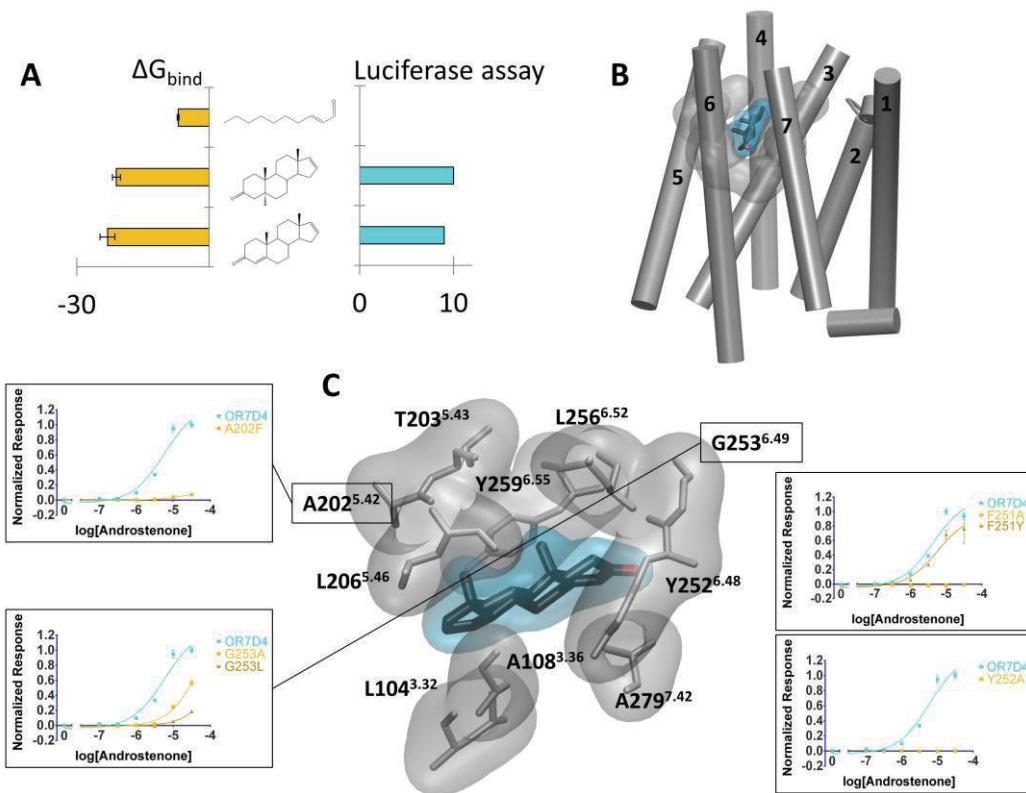


Figure 1. Binding cavity contribution to binding. A - Comparison of free energy calculation results (on the 3 MDs) with experimental data of *in vitro* response at 300 μ M.[10] B – homology model of hOR7D4. TMs are numbered from 1 to 7, the binding cavity is highlighted as a gray surface. C – Binding cavity of hOR7D4 bound to androstadienone (blue). Residues of the binding cavity are shown in gray. Dose-response data upon androstenone stimulation are shown for A202F, F251A, F251Y, G253A, and G253L mutant ORs.

Molecular dynamics discriminates agonists from neutral antagonist

Using a similar starting structure discussed above, a 3D atomistic model was built for each system (androstadienone, androstenone, and Z-(2)-decenal). Then, a set of three unconstrained molecular dynamics simulations was performed with the purpose of an *in silico* screening that would discriminate between agonists on one side, and the neutral antagonist on the other.

The model, in line with our previous findings on a constitutively active mutant model, is able to recover a structure typical of an active state for the two agonists.[9] When the ligand is a neutral antagonist, the models converge toward a structure typical of an inactive form. When androstenone and androstadienone are bound to hOR7D4, the molecular dynamics simulations rapidly converge towards structures showing the hallmarks of an activated form (Fig. 2). The intracellular part of TM6 shifts away from its initial position and gets far removed from the intracellular parts of TM3 and TM7. The ionic-lock between D3.50 and the backbone oxygen atom of R6.30 evolves from 10 Å to values close to 17 Å (Fig 2). In the case of the neutral antagonist (Z)-2-decenal, the system

reorganizes to sample a structure where the ionic-lock is closed with a distance $\sim 10 \text{ \AA}$ shorter than those found in the case of agonists (Fig. 2).

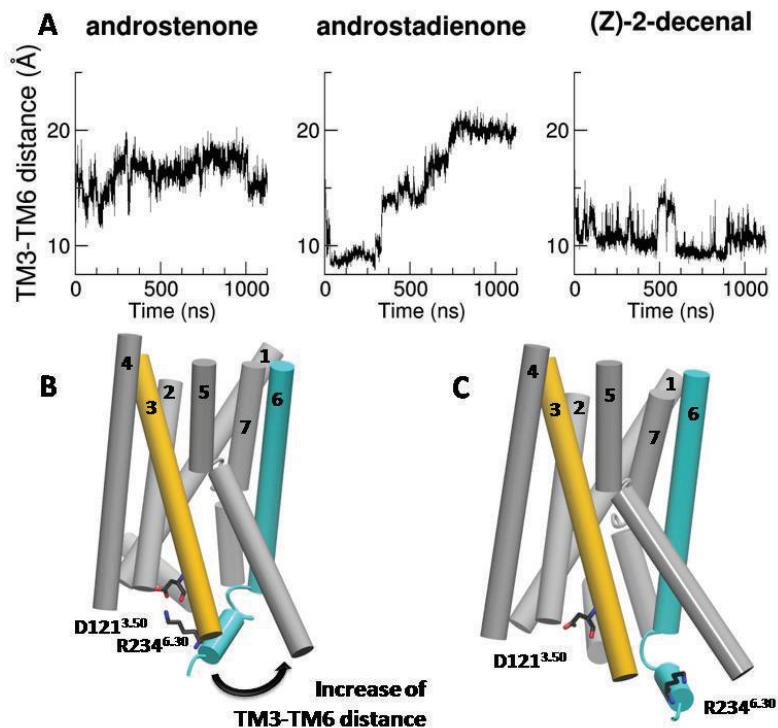


Figure 2. Active and inactive states of OR7D4 as observed by molecular dynamics simulations. A. Distance between D121 in TM3 and R234 in TM6 during one out of the three microsecond molecular dynamics simulations. B & C. Typical structures of inactive (B) and active states (C) of hOR7D4. TM3 is shown in yellow and TM6 is shown in cyan.

Molecular dynamics simulations are a computational microscope that predicts odorant receptor state upon odorant stimulation

In a previous article, we have shown that molecular modeling was able to predict the behavior of a receptor and some of its mutants associated to strong basal activities.[9] Here, we have extended the proof on concept of the predictive power of modeling by showing that this approach also works upon various odorants stimulations.

The model of hOR7D4 was built by homology and the same apo structure was bound to androstenone and androstadienone, two known agonists, and to Z-(2)-decenal, a known neutral antagonist *in vitro*. Upon a multiple molecular dynamics protocol, where each system is replicated twice, we have observed that in all of the three replicas, the system is evolving either towards structures showing the hallmarks of an active state in case of agonists binding, or towards an inactive structure when it is stimulated by the neutral antagonist. Site-directed mutagenesis has assessed the quality of the model by comparing the receptor response with those of some predicted mutants. We

have identified that in the highly specific and conserved motif within the TM6 ORs, (FYG), the two first residues must be aromatic while the third one should preferentially be tiny residues.

Above all, we have put forward the predictive power of molecular dynamics simulations for virtually screening a series of odorant compounds on a given mammalian odorant receptor. This approach will be of primary importance for selecting odorants to be tested prior to making experiments.

Methods and Materials

Site-Directed Mutagenesis

The coding sequence of OR7D4 was amplified from genomic DNA of C57BL/6 mice and subcloned into the pcDNA3.1/TOPO vector (Invitrogen) with an N-terminal tag of the first 20 amino acids of rhodopsin. Site-directed mutants were constructed using the Quikchange site-directed mutagenesis kit (Agilent Technologies). The sequences of all plasmid constructions were verified by both forward and reverse sequencing (DNA sequencing core facility, University of Pennsylvania).

Evaluation of OR surface expression

Live-cell immunostaining is used to evaluate OR surface expression.[12] Hana3A cells were co-transfected with the receptor and GFP plasmids 24 hours before the staining. The transfected Hana3A cells were incubated with the primary antibody solution (mouse anti-rhodopsin, Rho 4D2, Abcam) on ice for 1 h. After rinsing the cells for three times, the secondary antibody solution (Alexa Fluor 568-conjugated anti-mouse IgG) was added onto the cells, and incubated for 45 min on ice. At the end of the incubation, the cells were fixed with 2% Paraformaldehyde, and mounted with vectashield mounting medium (Vector Laboratories, Inc.). The ratio of Rho⁺ cells/GFP⁺ cells is used to evaluate the surface expression of each OR construct.

Luciferase assay in Hana3A cells

The Dual-Glo Luciferase Assay (Promega) was used to determine the activities of firefly and Renilla luciferase in Hana3A cells [12]. Firefly luciferase, driven by a cAMP response element promoter (CRE-Luc; Stratagene), was used to determine OR activation levels. Renilla luciferase, driven by a constitutively active SV40 promoter (pRLSV40; Promega), functioned as an internal control for transfection efficiency and cell viability. Hana3A cells stably expressing RTP1L, RTP2, REEP1, and G_{olf} were plated on poly-D-lysine-coated 96-well plates (Nalge Nunc) and incubated overnight in minimum essential medium eagle (Sigma) with 10% FBS at 37°C and 5%CO₂. The following day, cells were transfected using Lipofectamine 2000 (Invitrogen). For each 96-well plate, 1 µg pRL-SV40, 1 µg CRE-Luc, 1 µg mouse RTP1s, and 6 µg of receptor plasmid DNA were transfected. After transfection (24 h), medium was replaced with 25 µl of odorant solution diluted in CD293 chemically defined

Article 10 – de March et al. en préparation

medium (Invitrogen), and cells were further incubated for 4 h at 37°C and 5% CO₂. The manufacturer's protocols were followed to measure firefly luciferase and Renilla luciferase activities. A Wallac Victor 1420 plate reader (Perkin-Elmer) was used to measure luminescence. Data were analyzed using Microsoft Excel and GraphPad Prism. Normalized activity was further calculated using the following formula: [Luc/RLuc(N)-Luc/RLuc(lowest)]/[Luc/RLuc(highest)-Luc/RLuc(lowest)], where Luc/RLuc(N) = luminescence of firefly luciferase divided by luminescence of Renilla luciferase in a certain well; Luc/RLuc(lowest) = lowest firefly luminescence divided by Renilla luminescence of a plate or set of plates; Luc/RLuc(highest) = highest firefly luminescence divided by Renilla luminescence of a plate. To facilitate comparison between OR responses from multiple plates, the Rho-tag empty vector and wt OR7D4 were always included as negative and positive control, respectively. The basal activity of an OR was averaged from four wells in the absence of odorants and further corrected by subtracting that of the control empty vector. An odorant-induced activity was averaged from at least three wells and further corrected by subtracting the basal activity of that receptor. Both basal activity and odorant-induced responses were corrected for the surface expression ratio (Rho+/GFP+ when Hana3A cells were co-transfected with a Rho-tagged OR and GFP) normalized to that of wt.

Molecular modeling

Model building. The protocol follows a previously published method[13]. The sequence of OR7D4 is aligned with 395 human ORs[14] and nine sequences of X-ray elucidated GPCRs: bovine rhodopsin (PDB: 1U19)[15], human beta 2 adrenergic (PDB: 2RH1)[16], turkey beta 1 adrenergic (PDB: 2VT4)[17], human chemokine receptors CXCR4 (PDB: 3ODU)[18] and CXCR1 (PDB: 2LNL)[19], human dopamine receptor D3 (PDB: 3PBL)[20], human adenosine a2A receptor (PDB: 2YDV)[21], human histamine H1 receptor (PDB: 3RZE)[22] and muscarinic acetylcholine receptor M2 (PDB: 3UON)[23]. Highly conserved motifs in ORs are considered as constraints for the alignment: GN in helix 1, PMYFFLXXLSXXD in helix 2, MAYDRYXAICXPLXY in helix 3, SYXXI in helix 5, KAFSTCASH in helix 6, LNPXIY in helix 7 and a pair of conserved cysteines 973.25-1794.80 which constitute a known disulfide bridge between the beginning of helix 3 and the extracellular loop 2. Four experimental GPCR structures (1U19, 3ODU, 2YDV and 2LNL) are selected as templates to build OR7D4 by homology modeling with Modeller.[24] The N-terminal structure is omitted to avoid perturbing the modeling protocol. Five models are obtained and the one fulfilling several constraints (binding cavity sufficiently large, no large folded structure in extra-cellular loops, all TMs folded as α -helices, a small α -helix structure between TM3 and TM4) is kept for further molecular dynamics simulations.

Molecular dynamics simulations. The wt OR7D4 is embedded in a model membrane made-up of POPC lipids solvated by TIP3P water molecules using Maestro.[25] The total system is made up of ~48,650 atoms in a periodic box of 91*89*98 Å³.

Molecular dynamics simulations are performed with sander and pmemd.cuda modules of AMBER12 with the ff03 force-field for the protein and the gaff.lipid for the membrane. Hydrogen atoms bond are constrained by SHAKE algorithm and long-range electrostatics interactions are handled with Particle Mesh Ewald (PME). The cutoff for non-bonded interactions is set at 8 Å. Temperature and pressure are maintained constant with a Langevin thermostat with a collision frequency of 2 ps⁻¹. In addition, a weak coupling anisotropic algorithm with a relaxation time of 1 ps⁻¹ is applied. Snapshots are saved every 20 ps.

Two energy minimizations are performed during 10,000 steps with the 5,000 first steps using a conjugate gradient algorithm. The first one is run with a restraint of 200 kcal.mol⁻¹ applied on all atoms of the membrane and water and the second one with the same restraint on all atoms of the receptor. This last constraint is kept for the heating phase of 20 ps (NTP, 100K to 310K, Langevin thermostat with collision frequency of 5 ps⁻¹) and equilibration of 15 ns (NTP, 310K). Restraints are then reduced by 5 kcal.mol⁻¹Å⁻² and another cycle of minimization-equilibration is performed. The systems are replicated three times and 1200 ns-long production molecular dynamics are performed after an equilibration period of 50 ns.

References

- [1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: A molecular basis for odor recognition, *Cell* 65 (1991) 175-187.
- [2] B. Malnic, P.A. Godfrey, L.B. Buck, The human olfactory receptor gene family, *Proc Natl Acad Sci U S A* 101 (2004) 2584-2589.
- [3] Y. Yu, C.A. de March, M.J. Ni, K.A. Adipietro, J. Golebiowski, H. Matsunami, M. Ma, Broad Responsiveness of G Protein-Coupled Odorant Receptors Is Attributed to the Activation Mechanism, (submitted)
- [4] C.A. de March, S. Ryu, G. Sicard, C. Moon, J. Golebiowski, Structure–odour relationships reviewed in the postgenomic era, *Flavour and Fragrance Journal* 30 (2015) 331-410.
- [5] **Pharmacology of G Protein Coupled Receptors**, vol. 62: Academic Press; 2011.
- [6] R. Nygaard, Y. Zou, Ron O. Dror, Thomas J. Mildorf, Daniel H. Arlow, A. Manglik, Albert C. Pan, Corey W. Liu, Juan J. Fung, Michael P. Bokoch *et al*, The Dynamic Process of β2-Adrenergic Receptor Activation, *Cell* 152 (2013) 532-542.
- [7] A.J. Venkatakrishnan, X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, M.M. Babu, Molecular signatures of G-protein-coupled receptors, *Nature* 494 (2013) 185-194.
- [8] C.A. de March, J. Golebiowski, A computational microscope focused on the sense of smell, *Biochimie* 107 (2014) 3-10.

Article 10 – de March et al. en préparation

- [9] C.A. de March, Y. Yu, M.J. Ni, K.A. Adipietro, H. Matsunami, M. Ma, J. Golebiowski, Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors, *Journal of the American Chemical Society* 137 (2015) 8611-8616.
- [10] A. Keller, H.Y. Zhuang, Q.Y. Chi, L.B. Vosshall, H. Matsunami, Genetic variation in a human odorant receptor alters odour perception, *Nature* 449 (2007) 468-U466.
- [11] C.A. de March, S.-K. Kim, S. Antonczak, W.A. Goddard, J. Golebiowski, G protein-coupled odorant receptors: From sequence to structure, *Protein Sci* 24 (2015) 1543-1548.
- [12] H. Zhuang, H. Matsunami, Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells, *Nat Protoc* 3 (2008) 1402-1413.
- [13] L. Charlier, J. Topin, C.A. de March, P.C. Lai, C.J. Crasto, J. Golebiowski, in: C.J. Crasto (Eds.), *Olfactory Receptors, Molecular Modelling of Odorant/Olfactory Receptor Complexes*, New York, 2013, pp. 53-65.
- [14] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, *Genome Biol* 2 (2001) 1-12.
- [15] T. Okada, M. Sugihara, A.N. Bondar, M. Elstner, P. Entel, V. Buss, The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure, *J Mol Biol* 342 (2004) 571-583.
- [16] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka *et al*, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science* 318 (2007) 1258-1265.
- [17] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, G.F. Schertler, Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature* 454 (2008) 486-491.
- [18] B. Wu, E.Y.T. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi *et al*, Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, *Science* 330 (2010) 1066-1071.
- [19] S.H. Park, B.B. Das, F. Casagrande, Y. Tian, H.J. Nothnagel, M. Chu, H. Kiefer, K. Maier, A.A. De Angelis, F.M. Marassi *et al*, Structure of the chemokine receptor CXCR1 in phospholipid bilayers, *Nature* 491 (2012) 779-783.
- [20] E.Y. Chien, W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov *et al*, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, *Science* 330 (2010) 1091-1095.
- [21] G. Lebon, T. Warne, P.C. Edwards, K. Bennett, C.J. Langmead, A.G. Leslie, C.G. Tate, Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation, *Nature* 474 (2011) 521-525.
- [22] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han *et al*, Structure of the human histamine H1 receptor complex with doxepin, *Nature* 475 (2011) 65-70.
- [23] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga *et al*, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, *Nature* 482 (2012) 547-551.
- [24] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper, A. Sali, in: (Eds.), *Current Protocols in Bioinformatics, Comparative Protein Structure Modeling Using Modeller*, 2006, pp.
- [25] Schrödinger: **1: Maestro, version 9.4.** In: *Schrödinger, LLC, New York, NY*. Edited by L. Schrödinger, New York, NY; 2013.

Conclusions et perspectives

Les objectifs de cette thèse de doctorat étaient d'établir les bases fondamentales nécessaires à l'étude des relations entre la structure des molécules odorantes et leur odeur sous un nouvel angle. Cette approche originale consiste à prendre en compte les protagonistes biologiques impliqués dans la physiologie de l'olfaction des mammifères. Les ROs sont la pierre angulaire de notre système olfactif. Ils transforment le signal chimique que sont les molécules odorantes, en signal électrique se propageant dans notre cerveau et qui est interprété comme une odeur.

Il a été nécessaire de partir des connaissances actuelles disponibles sur nos ROs. Une étude bibliographique des liens entre ROs et molécules odorantes a été réalisée et a permis d'identifier les couples ROs-agonistes connus. Cette base de données, issues de résultats expérimentaux, a été le socle de la recherche que nous avons menée par la suite. L'amplitude des espaces ROs (environ 400) et des molécules odorantes (virtuellement infini) rend l'investigation expérimentale et exhaustive des couples RO-odorant impossible. La mise en place d'un criblage virtuel apparaît comme une alternative aux approches expérimentales. Notre objectif était d'obtenir un outil informatique capable de recréer un nez virtuel physiologiquement pertinent en modélisant les ROs et leur comportement face à un odorant. Une série de questions, partant des relations structure-odeur jusqu'à la compréhension des mécanismes impliqués dans l'olfaction des mammifères, s'est posée :

- Peut-on trouver des caractéristiques physico-chimiques communes à des molécules d'une même famille olfactive malgré leur variabilité structurale ?
- Ces modèles physico-chimiques, uniquement basés sur des structures de molécules odorantes, peuvent-ils être reliés au code combinatoire de récepteurs activés ?
- Dans quelle mesure le calcul des constantes thermodynamiques d'un couple RO-odorant permet-il de prédire l'activation d'un récepteur par une molécule odorante ?
- Quels sont les mécanismes régulant le spectre de reconnaissance d'un RO ?
- L'activation d'un neurone, *via* l'activation du RO qu'il exprime, peut-elle être prédictive grâce à la modélisation moléculaire ? De manière plus précise, la détermination du caractère agoniste ou non-agoniste d'une molécule peut-elle se faire par l'observation *in silico* de l'activation d'un RO ?

Ces questions ont guidé des études au laboratoire, réalisées en collaboration avec des chimistes, biologistes et neuroscientifiques. Nous avons fait le choix de commencer par la réalisation d'un olfactophore, qui est à ce jour une méthode couramment appliquée à l'étude des relations structure-odeur. Elle consiste à définir les points communs structuraux entre les molécules odorantes d'une même famille olfactive. L'olfactophore de l'odeur santalée a été créé avec l'aide de nos collaborateurs en chimie de synthèse. Afin de contourner la grande variabilité structurale au sein des molécules à odeur santalée, le groupe d'odorants a été divisé en trois grandes familles de structures. Les modèles de l'olfactophore ainsi générés sont capables de discriminer les molécules santalées des molécules non-santalées. Cette méthode permet au chimiste de faire un premier tri dans une série de molécules potentiellement odorantes et de n'envisager la synthèse que des odorants prédis par le modèle comme appartenant à la famille olfactive voulue. Toutefois, les mécanismes de cette discrimination restent obscurs. Ces modèles sont uniquement basés sur des structures de composés odorants en omettant de prendre en compte la physiologie de notre olfaction. Dans quelle mesure ces résultats pourraient être utiles à l'identification de récepteurs impliqués dans le code combinatoire de cette odeur ?

La superposition de nos modèles d'olfactophore avec les pharmacophores de récepteurs olfactifs connus a révélé une grande similarité entre l'une de nos hypothèses de l'odeur santalée et le pharmacophore du récepteur hOR1G1. Cette constatation est cohérente avec le fait que hOR1G1 est identifié comme répondant positivement au camphre ou aux odorants à longue chaîne carbonée, qui présentent tous les deux des similarités avec les molécules santalées. Les expériences obtenues *in vitro* ont confirmé nos hypothèses. Les caractéristiques du récepteur hOR1G1 étaient encodées dans l'olfactophore des molécules de la famille des santalanes. Cette découverte ouvre la voie à la revisite des modèles d'olfactophores existants pour tenter d'identifier des ROs impliqués dans le code combinatoire de l'odeur associée.

La suite des travaux s'est concentrée sur la compréhension des mécanismes d'activation des récepteurs olfactifs. Là encore, les approches de modélisation moléculaire sont couplées à des données expérimentales. Dans le but de créer un système de criblage virtuel permettant de prédire les molécules associées à un récepteur, la performance du calcul d'affinité ligand-récepteur a été évaluée. Nous avons montré que cette grandeur thermodynamique permet de discriminer de façon efficace les molécules agonistes des non-agonistes de hOR1G1 sur un groupe de 10 composés de familles chimiques variées. Les capacités de cet outil peuvent donc être utilisées pour mieux comprendre le spectre de reconnaissance des ROs, tenter de prédire les affinités de molécules odorantes pour un récepteur et, à terme, déchiffrer le code combinatoire lié aux odeurs.

Les récepteurs olfactifs possèdent, pour une grande part d'entre-eux, la particularité de répondre différemment à des composés odorants de la même famille chimique. La subtilité de leur spectre de reconnaissance peut être regroupée en deux grands comportements : les ROs à large spectre de reconnaissance et ceux à spectre restreint. Les caractéristiques qui dictent ces comportements étaient encore mal comprises.

La combinaison de l'analyse de séquence de récepteurs de la souris, de leur étude par modélisation moléculaire et de validation par mutagénèse dirigée nous a permis de déterminer que le spectre de reconnaissance d'un RO était régulé partiellement par la permissivité de sa cavité de liaison aux odorants mais aussi par sa capacité à s'activer. Cette information nous permet de mieux comprendre le fonctionnement de nos ROs et souligne l'importance du phénomène d'activation du récepteur dans la discrimination d'odorants. Connaitre l'affinité d'un ligand pour la cavité d'un récepteur est utile mais la subtilité des mécanismes d'activation des récepteurs rend cette information insuffisante pour déterminer si cette molécule est capable d'activer un neurone olfactif.

Nous nous sommes focalisés sur les mécanismes d'activation des ROs. Il n'existe aucune structure expérimentale de ces récepteurs. Grâce à la modélisation moléculaire et à des analyses bio-informatiques couplées à des validations expérimentales, des résidus conservés chez les ROs ont été identifiés comme contrôlant leur activation. Les simulations de dynamique moléculaire sont capables de prédire des phénomènes d'activation chez un RO de la souris, mOR256-3 et humain, OR7D4. Ces évènements théoriques sont en totale concordance avec les données expérimentales *in vitro* obtenues par nos collaborateurs. Ils nous ont permis de déterminer le réseau d'acides aminés cruciaux pour l'activation des ROs. Nous avons montré pour la première fois que la modélisation moléculaire était capable de prédire l'activation d'un RO (donc d'un neurone olfactif) lorsqu'il est stimulé par un candidat agoniste. Cette méthode établit des bases scientifiques robustes pour tenter d'élucider le code combinatoire d'activation de RO associé aux odeurs.

Grâce à ce travail de thèse, les outils fondamentaux permettant d'envisager l'établissement de relations structure-odeur sur une base physiologiquement inspirée semblent identifiés. Un protocole robuste permettant de construire les structures de l'ensemble des ROs de mammifères est désormais disponible. Aussi, le criblage virtuel est désormais accessible à travers les protocoles mis au point et testés.

Il restera dans le futur à réaliser un travail récapitulatif afin de tester l'hypothèse de code combinatoire pour une famille odorante donnée. Il s'agit de savoir si nous sommes en mesure de reproduire un code combinatoire d'activation expérimentale de ROs pour une molécule odorante. En

extrapolant ce travail à un groupe de composés d'une même famille olfactive, nous espérons déchiffrer le code combinatoire de cette odeur et être ainsi capable de générer de nouvelles structures de molécules possédant cette odeur. Ainsi, un « nez virtuel » est sur le point de voir le jour.

Au-delà de l'olfaction, les ROs sont exprimés dans différents organes de notre corps comme les reins, l'intestin, le cerveau, l'estomac, les testicules ou certaines cellules cancéreuses. Les résultats de nos recherches vont trouver leur application dans bien d'autres domaines que notre perception des odeurs. Par exemple, les ROs sont en partie associés à la chimiotaxie des spermatozoïdes ou sont suspectés d'être impliqués dans les phénomènes de satiété. De plus, la découverte des PSGRs (Prostate-Specific Gene Receptors), dont les gènes sont homologues à ceux des ROs, dans les cellules de la prostate et surexprimés en cas de cancer, fait des ROs une nouvelle cible thérapeutique. La compréhension des mécanismes de blocage/activation de ces récepteurs est donc d'un intérêt fondamental dans des domaines de recherche aussi variés que le contrôle de la fécondation, la diététique ou le traitement contre le cancer.

Annexe

Annexe

La mécanique moléculaire

Pour des systèmes de taille importante ou lorsqu'une étude *dynamique* d'un système moléculaire est envisagée, l'utilisation de méthodes de la chimie quantique peut s'avérer trop coûteuse en temps de calculs. Les méthodes appelées du *champ de force*, considèrent l'énergie électronique comme un paramètre dépendant des coordonnées nucléaires. Ces paramètres sont optimisés sur des données expérimentales ou des calculs de chimie quantique. Les molécules sont alors représentées comme des groupements d'atomes liés entre eux par des liaisons. Ces atomes peuvent avoir différentes tailles et duretés, les liaisons quant à elles, étant plus ou moins souples. Ces méthodes du champ de force sont alors référencées comme des méthodes de *mécanique moléculaire (MM)*.

L'énergie du champ de force résulte de différentes contributions à l'énergie totale, on l'écrit :

$$E_{Champ de Force} = E_{liaisons} + E_{angles} + E_{torsion} + E_{vdW} + E_{électrostatique} + (E_{croisés})$$

où chaque terme correspond à la déformation des liaisons, des angles des angles dièdres, de l'éloignement des atomes non liés ou de termes croisés entre ces différentes contributions. Les conformations stables de systèmes moléculaires correspondent aux minimums de cette fonction énergie, elle-même dépendante des coordonnées nucléaires. La forme la plus simple, ne contenant que des potentiels harmoniques, est typiquement celle du champ de force AMBER :

$$\begin{aligned} E_{pot} = & \sum_{liaisons} k_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihèdres} \frac{V_n}{2}(1 + \cos(n\varphi - \gamma)) + \\ & \sum_{non_liés} \left[\frac{q_i q_j}{r_{ij}} + \epsilon_{ij} \left(\left(\frac{R_{ij}^*}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^*}{r_{ij}} \right)^6 \right) \right] \end{aligned}$$

Le faible coût de ces méthodes permet de réaliser un nombre important de calculs portant sur différentes conformations moléculaires. L'application de la seconde loi de Newton permet de simuler le mouvement des atomes. L'énergie est constituée de l'énergie potentielle et cinétique du système

considéré, laissant ainsi la possibilité aux molécules de surmonter les barrières énergétiques séparant les minimums de la surface d'énergie, on parle de *dynamique moléculaire*.

La dynamique moléculaire

Rappelons la seconde loi de Newton qui relie la force agissant sur une particule à sa masse et à son accélération :

$$F = m\mathbf{a}$$

que l'on peut aussi écrire sous la forme :

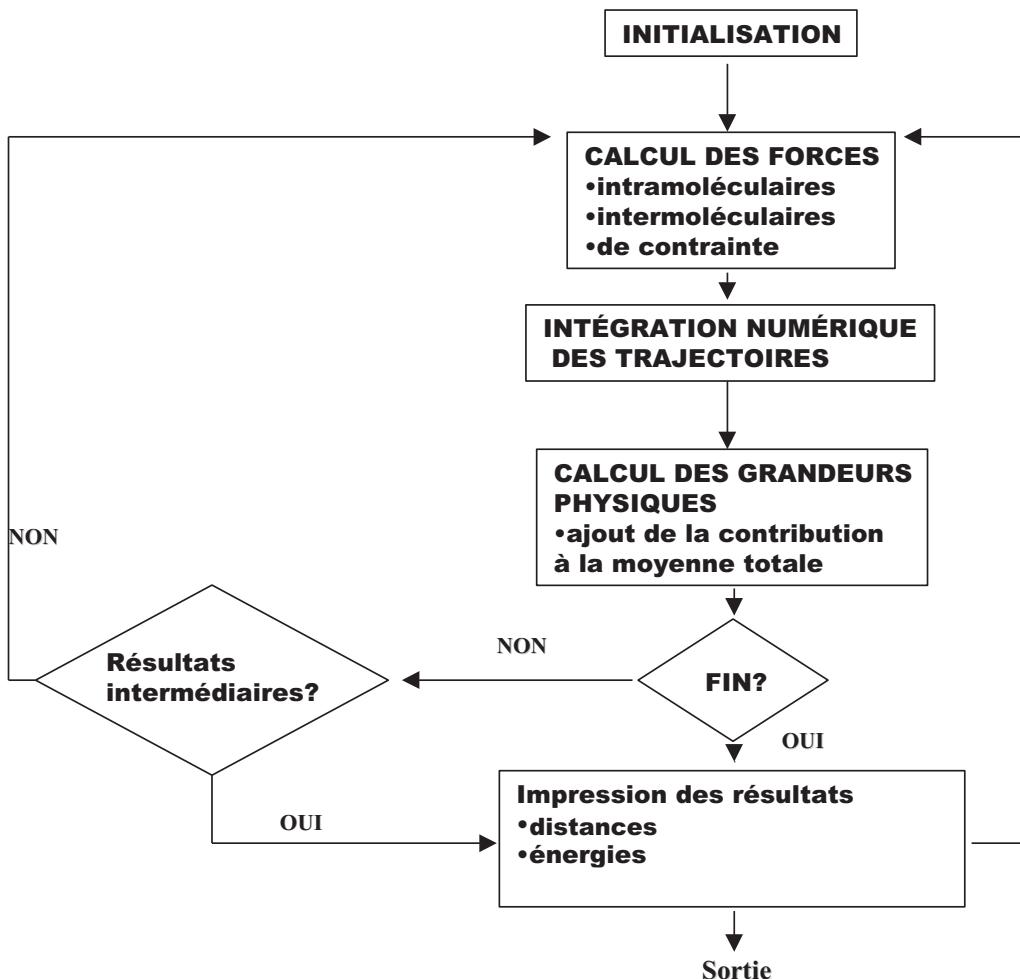
$$-\frac{dE_{pot}}{d\mathbf{r}} = m \frac{d^2\mathbf{r}}{dt^2}$$

où E_{pot} est l'énergie potentielle à la position \mathbf{r} . Le vecteur \mathbf{r} contient les coordonnées de toutes les particules du système.

Pour un jeu de particules aux positions \mathbf{r}_i , les positions à un temps $t+\Delta t$ sont obtenues par extension en une série de Taylor impliquant la vitesse et l'accélération des particules.

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_i \Delta t + \frac{1}{2} \mathbf{a}_i (\Delta t)^2 + \dots$$

Le pas d'intégration Δt est considéré en fonction du mouvement moléculaire le plus rapide, le plus souvent, il est de l'ordre de 10^{-15} secondes. Pour les systèmes moléculaires courants, le mouvement le plus rapide est l'élongation des liaisons impliquant un atome d'hydrogène. Ce degré de liberté n'influence que très peu voir pas du tout les propriétés calculées dans les dynamiques moléculaires. Il peut donc être intéressant de fixer ces liaisons à une distance donnée, afin de pouvoir augmenter le pas d'intégration Δt .



Principe général d'une simulation de Dynamique Moléculaire

Les conditions périodiques

L'étude de systèmes en phase condensée, implique la prise en compte d'un grand nombre de particules autour du site d'intérêt. Cependant, pour des raisons de réduction de temps de calcul, ce nombre doit être le plus restreint possible, sans compromettre la précision du résultat.

Afin de réduire le nombre de molécules étudiées, tout en simulant un système le plus proche possible d'une phase condensée, la boîte contenant le système moléculaire est répliquée dans les trois directions de l'espace, formant ainsi un réseau infini. Si une particule quitte la boîte « par le haut », elle est immédiatement introduite par le bas. De plus, les interactions entre particules sont tronquées à une distance étant égale au maximum à la moitié de la taille de la boîte, afin qu'une particule n'interagisse jamais avec elle-même, on parle alors de distance de *cut-off*.

La méthode MM-GBSA

Cette méthode est dite "end-point" car seules les énergies des états initiaux et finaux sont pris en compte. Elle est basée sur l'analyse de configurations extraites de simulations de dynamique moléculaire à l'équilibre. L'enthalpie libre est alors estimée suivant l'équation :

$$G = E_{MM} + H_{rot/trans} + G_{sol} - TS$$

où E_{MM} correspond à l'énergie potentielle, $H_{rot/trans}$ aux 6 degrés de liberté de rotation et translation ($H_{rot/trans} = 6 * \frac{1}{2}RT = 1,8 \text{ kcal.mol}^{-1}$ à 300 K) et S à l'entropie calculée grâce aux modes normaux.

G_{sol} est l'enthalpie libre de solvatation qui est estimée de façon identique au cas de la solvatation implicite.

La différence d'enthalpie libre est ensuite obtenue en calculant l'enthalpie libre pour chaque composant (ligand, récepteur, complexe) :

$$\Delta G = \langle G_{complexe} \rangle - \langle G_{ligand} \rangle - \langle G_{récepteur} \rangle$$

où $\langle G_x \rangle$ correspond à la moyenne de l'enthalpie libre du système X lors de la simulation de dynamique moléculaire.

Le docking

Le docking permet de générer des complexes ligands/récepteurs dont les conformations sont jugées suivant une fonction de scoring. Cette fonction de scoring, pour la plupart des algorithmes, est approximée à l'enthalpie libre de liaison. Pour Vina, le logiciel de docking le plus utilisé, cette fonction est basée sur le principe des relations structure-activité : l'énergie issue de la conformation du ligand est directement corrélée à son activité. L'enthalpie libre de liaison est alors approximée à partir de potentiels tirés de données expérimentales (knowledge-based).